A NETWORK EXTENDED-DIRECTIONAL MIX-EFFICIENCY MEASURE IN THE PRESENCE OF UNCONTROLLABLE INPUT AND UNDESIRABLE OUTPUT

BY

John Ogabor USHIE
B.Tech. (Yola), M.Sc. (Ibadan)
PG/PSC2216041

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BENIN,
BENIN CITY

November, 2025

A NETWORK EXTENDED-DIRECTIONAL MIX-EFFICIENCY MEASURE IN THE PRESENCE OF UNCONTROLLABLE INPUT AND UNDESIRABLE OUTPUT

BY

John Ogabor USHIE
B.Tech. (Yola), M.Sc. (Ibadan)
PG/PSC2216041

A THESIS WRITTEN IN THE DEPARTMENT OF MATHEMATICS AND SUBMITTED TO THE COLLEGE OF POSTGRADUATE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF DEGREE OF M.PHIL IN INDUSTRIAL MATHEMATICS OF THE UNIVERSITY OF BENIN, BENIN CITY, EDO STATE, NIGERIA.

November, 2025

CERTIFICATION

This is to certify that this thesis was written by John Ogabor USHIE with Matriculation Number PG/PSC2216041 in the Department of Mathematics, Faculty of Physical Sciences, University of Benin, Benin City, Nigeria under the supervision of Professor A. A. OSAGIEDE, Co-supervisor Dr. C. B. IBE.

Sign & Date ________________________

Professor A. A. OSAGIEDE
(Supervisor)

Sign & Date ________________________

Dr. C. B. IBE
(Co-supervisor)

Sign & Date ________________________

Professor D.U. OKUONGHAE
(Head of Department)

DECLARATION

UNIVERSITY OF BENIN
SCHOOL OF POSTGRADUATE STUDIES
A NETWORK EXTENDED-DIRECTIONAL MIX-EFFICIENCY MEASURE IN THE PRESENCE OF UNCONTROLLABLE INPUT AND UNDESIRABLE OUTPUT

BY

John Ogabor USHIE
B.Tech. (Yola), M.Sc. (Ibadan)
PG/PSC2216041

THE BOARD OF EXAMINERS DECLARE AS FOLLOWS, THAT THIS IS THE ORIGINAL WORK OF THE CANDIDATE, THAT THE THESIS IS ACCEPTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF MASTER OF PHILOSOPHY (M. Phil) IN INDUSTRIAL MATHEMATICS.

______________________-

Professor A. A. OSAGIEDE
Chief Supervisor

____

Dr. C. B. IBE
Co- supervisor

______________________________-

Professor D.U. OKUONGHAE
Head of Department

DECLARATION

UNIVERSITY OF BENIN

COLLEGE OF POSTGRADUATE STUDIES

A NETWORK EXTENDED-DIRECTIONAL MIX-EFFICIENCY MEASURE IN THE PRESENCE OF UNCONTROLLABLE INPUT AND UNDESIRABLE OUTPUT

BY

John Ogabor USHIE
B.Tech. (Yola), M.Sc. (Ibadan)
PG/PSC2216041

THE BOARD OF EXAMINERS DECLARES AS FOLLOWS, THAT THIS IS THE ORIGINAL WORK OF THE CANDIDATE, THAT THE THESIS IS ACCEPTED IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE AWARD OF THE DEGREE OF MASTER OF PHILOSOPHY (M.Phil.) IN MATHEMATICS.
I hereby grant the University of Benin, through the University Library, a non-exclusive, worldwide right to reproduce and distribute my Thesis and abstract (herein after ”the Work”), in whole or in part, through any media, in its present form or any translated version for preservation and accessibility, provided such translation does not alter its content. This grant is royalty-free, and I retain the right to publish the Work in its current or future version elsewhere.

Warranties

I further affirm that:

1. I am the sole author of the Work and grant the University of Benin the right to make it available four (4) years after the award of my M. phil degree, in compliance with the University of Benin Senate regulations.
2. The Work does not contain confidential information requiring third-party consent for disclosure.
3. I have exercised due diligence to ensure that the Work is original and does not breach any Nigerian law or infringe upon any third party’s copyright or other Intellectual Property Rights, to the best of my knowledge.
4. Where the Work includes copyrighted material not owned by me, I have obtained unrestricted permission from copyright holder to grant this license to the University of Benin Library. Such third-party materials are clearly identified and acknowledged within the Work.
5. In the event of any copyright dispute concerning the Work, I agree to indemnify and hold harmless the University of Benin, its officers, employees, and agents from any liability arising from the material authorized under this agreement.
6. The University of Benin is under no obligation whatsoever to take legal action on my behalf as the Depositor in the event of an intellectual property rights infringement or any other related dispute in the material deposited.

John Ogabor USHIE Date Author/Student

NAME	SIGNATURE	DATE
SUPERVISOR	....	....
CO-SUPERVISOR	....	....
HEAD OF DEPARTMENT	....	....

CERTIFICATION OF THESIS ON PLAGIARISM

We the undersigned attest and declare that the thesis of John Ogabor USHIE Titles: A Network Extended-Directional Mix-Efficiency Measure in the Presence of Uncontrollable Input and Undesirable Output has successfully passed the anti-plagiarism test and does not violate any copy right regulation.

____________________ ____________________

Professor A. A. OSAGIEDE **Dr. C. B. IBE**

Chief Supervisor Co- supervisor

____

Professor D.U. OKUONGHAE
Head of Department

DEDICATION

This project is dedicated to Almighty God.

ACKNOWLEDGMENT

I sincerely thank God for seeing me through this research work. My thanks also go to my supervisors, Professor A. A. OSAGIEDE and Dr. C. B. IBE, my mentors, who demonstrated fatherly role in addition, for their academic and professional guidance necessary to make this work come to acceptable completion. May God bless you and elongate your life in Jesus mighty name, Amen.

I am equally grateful to my Head of Department, Professor D. U. Okuonghae for his advise and encouragement. I extend my appreciation to the PG Coordinator Dr. I. I. Ako for your continuous encourgement, Professor C. I. Nkeki for your assistance and willingness to support, Dr. Mrs. R. Omoregie for accepting me and your encouragement, and all Postgraduate Board members and lecturers, I am grateful for your efforts in making this M. Phil. programme come to a successful ending.

I am also grateful to Mrs. R. Adewole for her support. God bless you all. I am indeed very grateful to my dear wife Mrs Monica Ushie for the encouragement, care, love, support and prayers toward the success of this programme. To my lovely children, Ukanada, ukwunada and Irinen, thank you so much for your prayers and understanding.

I also appreciate my siblings for standing by me. My Father in-law and mother in-law have been wonderful, thank you friends and well wishers for your support and prayers which have been major factors to the success of this research work. May God bless and preserve your life in Jesus name!

Table of Contents

Table of Figures

Table of Tables

ABSTRACT

In conventional data envelopment analysis, decision making units are considered as a whole, that is, internal stages are generally ignored. In real life, most systems are composed of many divisions operating interdependently via the intermediate products that are created by some divisions and consumed by some others within the system. The aim of the study was to develop a formalized network mix efficiency model that can deal with both controllable and uncontrollable inputs as well as desirable and undesirable outputs by considering the internal structures of systems.

A linear programming approach of data envelopment analysis was used by taking the ratio of slack based measure and directional distance function to measure the eco-efficiency of 54 African countries. The inputs into the system were classified as controllable and uncontrollable while the output as desirable and undesirable. Stage one captured agricultural production efficiency, while stage two evaluated the environmental impact of carrying out agricultural activities.

The network efficiency model identified sources of inefficiencies from the components of a complex system rather than looking at the system as a black box. Controllable inputs and undesirable outputs were minimized while the desirable outputs were maximized. The approach provided policymakers with robust benchmarks for enhancing agricultural productivity while mitigating environmental degradation.

CHAPTER ONE

Introduction

Data Envelopment Analysis (DEA) is a methodology based on the application of linear programming approach for evaluating the performance of a set of peer entities called Decision-Making Units (DMUs), which convert multiple inputs into multiple outputs. Data envelopment analysis have seen a great variety of applications in evaluating the performances of different kind of entities engaged in many different activities in many different contexts in many different countries. These DEA applications have used DMUs of various forms to evaluate the performance of entities, such as hospitals, US Air Force wings, universities, cities, courts, business firms, and others, including the performance of countries, regions, etc. Because it requires very few assumptions, DEA has also opened up possibilities for use in cases that have been resistant to other approaches because of the complex (often unknown) nature of the relations between the multiple inputs and multiple outputs involved in DMUs. DEA evaluates the relative efficiency of a peer set of entities called DMUs, which consume multiple inputs to produce multiple outputs. In the field of performance measurement, DEA has emerged as a reliable technique for assessing efficiency and establishing targets by identifying benchmarks. DEA is categorized into two main approaches: black-box and network DEA.

Black-box DEA utilizes initial inputs to generate final outputs without looking into internal processes, thereby overlooking inefficiencies.

However, the network DEA considers internal processes to identify the sources of inefficiency for each DMU.

The introduction of a network structure within DEA is to address the limitations of the black-box DEA model. This enhanced model, known as network DEA (NDEA), incorporates network structures to capture more refined efficiency considerations. The network structure can manifest in various forms, including series, parallel and mix structures. In a two-stage DEA model, the first stage employs inputs to generate intermediate measures, which then serve as inputs for the second stage to produce final outputs.

The Directional Distance Function (DDF) model is a radial input-oriented and output-oriented models proposed by Chambers (1996) to measure a set of n Decision Making Units $(DMUs_j$, j = 1, ..., n). The DDF model measures the distance from a particular combination of inputs-output (x, y)$\in \Re^{m+s}$ to the efficiency frontier of the technology set T in the direction vector.

The Slack Based Measure (SBM) is a non oriented DEA model that considers both input excesses and output shortfalls simultaneously while dealing with their slacks. it project an efficient DMU onto the efficient frontier,Tone (2001) . More so, the SBM efficiency score leaves no output or input un-calculated because the target function takes into account all potential improvements to outputs and inputs. The SBM evaluates the efficiency of DMU_j (j = 1, ..., n) by solving the fractional programme.

1.1 Definitions

Definition 1.1. *Data Envelopment Analysis(DEA) is a methodology based upon the application of linear programming approach for evaluating the performance of a set of peer entities called decision making units (DMU), which converts multiple inputs into multiple outputs.*

In general terms, the essential idea is that we wish to assess how efficiently each decision making unit is handling the transformation process when compared with other decision making units engaged in the same process.

Definition 1.2. *Decision Making Units (DMU) refer to any entity that is to be evaluated in terms of its ability to convert inputs to outputs.*

Definition 1.3. *Efficiency (Extended Pareto - Koopmans Definition): full (100%) efficiency is attained by any decision making units if and only if non of it input or output can be improved without worsening some of its other input or output.*

Definition 1.4. *DEA Efficiency: the performance of DMU_0 is fully (100%) efficient if and only if both*

(1) $\theta^* = 1$ *and,*

(2) all slack $S_i^{-*} = S_r^{+*} = 0$.

Definition 1.5. *Weakly DEA Efficiency: the performance of* DMU_0 *is weakly efficient if and only if both*

(1) $\theta^* = 1$ *and,*

(2) $S_i^{-*} \neq 0$ *and/or* $S_r^{+*} \neq 0$*, for some i or r in some alternate optima.*

Definition 1.6. *Relative Efficiency : a DMU is said to be rated as fully (100%) efficient on the basis of available evidence if and only if the performance of other DMUs does not show that some of it input or output can be improved without worsening some of its other inputs or outputs.*

Pareto Efficiency: depending on whether inputs or outputs controllable, different measures of efficiency are appropriate. Two definitions are given, the one labeled output orientations is appropriate when outputs are controllable and the one labeled input orientation is appropriate when inputs are controllable. Output Orientation: a DMU is Pareto- efficient if it is not possible to raise any one of its output levels without lowering at least another one of its output level and/or without increasing at least one of its input.

Input Orientation: a DMU is Pareto- efficient if it is not possible to lower any one of its input levels without increasing at least another one of its input level and/or without lowering at least one of its output levels.

Mathematically, let $y_{rj}(r = 1, ..., s)$ be the output levels secured by DMU_j and x_{ij} the levels of input (i = 1, ..., m) it uses.

Slack Based Measure of Efficiency is formulated as

$$(SBM_{in}) \min \; p_{in} = 1 - \frac{1}{m}\sum_{i=1}^{m}\frac{S_i^-}{x_{io}} \qquad (1.1.1)$$

subject to :

$x_o = X\lambda + S^-,$

$y_o = Y\lambda - S^+,$

$\lambda \geq 0, S^- \geq 0, S^+ \geq 0,$

where s^- and s^+ and the slack variables for input and output respectively.

let an optimal solution of (SBM_{in}) be $(P_{in}, \lambda^*, S^{-*}, S^{+*})$ then we have the relationship $P_{in} \leq \theta^*_{CCR}$.

Definition 1.7. *Mix Efficiency: let the input-oriented CCR and SBM score of DMU_0 be θ^*_{CCR} and P^*_{in} respectively. the mix efficiency is defined by*

$$Mix = \frac{P^*_{in}}{\theta^*_{CCR}}.$$

Definition 1.8. *A system is said to be a **network** if there exist internal structures link to the working of the system as a whole, the output from one stage will become an input to the next stage.*

Definition 1.9. ***Controllable input**: an input is said to be controllable if it is at the discretion or under the control of the decision maker.*

Definition 1.10. ***uncontrollable input** are input that are not under the discretion or the control of the decision maker.*

Definition 1.11. ***Desirable output** are good output desired by the decision maker.*

Definition 1.12. ***Undesirable output** are bad output that are not desired by the decision maker.*

Definition 1.13. *A convex combination is a way to create a new point from a set of points. Given points $x_1, x_2, ..., x_n$, a convex combination is:*

$\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_n x_n$,

where,

$\lambda_i \geq 0 \quad \forall i, \sum_{i=1}^{n} \lambda_i = 1.$

1.2 Statement of the problem

In conventional DEA, DMUs are considered as a black box (whole), in that component structures are generally ignored. Most systems are composed of many divisions operating interdependently via the intermediate products that are created by some divisions and consumed by some others within the system. Ignoring the operations of the component divisions may cause misleading results in efficiency evaluation in the presence of uncontrollable input and undesirable output. Consequently, the main questions to be addressed in this research includes:

1. how can uncontrollable inputs and undesirable outputs be modelled into non-radial DDF and SBM network models?
2. how can efficiencies of component divisions and the whole system be evaluated simultaneously?

1.3 Aim and Objectives

The aim of the study was to develop a formalized network mix efficiency model that can deal with both controllable and uncontrollable inputs as well as desirable and undesirable outputs by considering the internal structures of systems.
The objectives of the study were to:

1. determine the efficiencies of internal stages and the whole system; and
2. use a comprehensive quantitative approach to measure system efficiency by considering both controllable and uncontrollable inputs as well as desirable and undesirable outputs.

1.4 Justification and Motivation for the Study

The main objective is to provide a rigorous qualitative and quantitative study of a network extended-directional mix-efficiency measure in the presence of controllable/uncontrollable input and desirable/undesirable output. Also, to gain deeper insight into the dynamics of data envelopment analysis. In the same vein, this thesis is centered on developing new, strong, and realistic models that can consider the efficiency measure of the components of a complex system and the system as a whole, providing detailed qualitative and quantitative analyses with emphasis on determining the existence of inefficiencies in the internal structures of a system. The knowledge of identifying inefficiency is very important in determining what is to be improved upon to make the system fully efficient. This also helps to determine the position of decision making units and rank them accordingly.

1.5 Thesis Structure

This thesis is organised as follows: chapter 2 reviews the literature; chapter 3 presents methodology; chapter 4 shows results and discussion; chapter 5 summary and conclusion.

CHAPTER TWO

Literature Review

2.1 Introduction

Depending on the available data, the production process under examination, and the research question, specific classes of models are being used. A first distinction is between parametric and non-parametric models.

2.1.1 Parametric models

a priori (using facts or principles that are known to be true in order to decide what the probable effects or results of something will be) assumptions about the production function are made. Stochastic Frontier Analysis (SFA) is the most representative of the parametric models.

Advantages

1. Statistical inference is well established as the noise is explicitly accounted for, while in the non-parametric case all the deviations from the frontier are interpreted as ineffiiency.

2. it is easier to account for environmental factors and endogeneity.

Limitations

1. They are not fully multidimensional since they are limited to one input or one output.

2. Shaping the characteristics of models with multiple inputs and outputs become immediately complex (e.g. for the scale efficiency).

2.1.2 Semi-parametric models

As a middle ground between parametric and non-parametric models, one can rely on the semi-parametric models. In this case, the econometric and the axiomatic approaches to efficiency analysis are integrated; the set of assumptions is restricted with respect to the parametric case, but the form of the production function is not fully specified, and the error terms is still distinguished from the inefficiency terms. As such, these models combine the virtue of SFA and DEA. Semi-parametric models have been fine-tuned to adapt to the panel structure of data.

Advantages

1. It improve statistical efficiency.

2. A balance between parametric and non-parametric models.

Limitations

1. the semi-parametric approaches are in general computationally heavier than standard parametric model.

2. prone to the curse of dimensionality.

2.1.3 Non-parametric Models

Charnes et al. (1978), formalize the maximization problem proposed by Farrell (1957) in a multiple-input multiple-output case by using linear programming which represents a constant returns to scale DEA model.

They assume that there are n DMUs to be evaluated. Each DMU consumes varying amounts of m different inputs to produce s different outputs. Specifically, DMU_j consumes amount x_{ij} of input i and produces amount y_{rj} of output r. They assume that $x_{ij} \geq 0$ and $y_{rj} \geq 0$ and further assume that each DMU has at least one positive input and one positive output value. We now turn to the "ratio-form" of DEA. In this form, as introduced by Charnes et al. (1978), the ratio of outputs to inputs is used to measure the relative efficiency of the $DMU_j = DMU_o$ to be evaluated relative to the ratios of all of the j = 1, 2, . . ., n.

$$Max \quad h_o(u,v) = \frac{\sum_r u_r y_{ro}}{\sum_i v_i x_{io}}, \tag{2.1.1}$$

where u, v are weight associated with output and input respectively. (2.1.1) is unbounded. Normalization constraint was introduced.

$$Max \quad h_o(u,v) = \frac{\sum_r u_r y_{ro}}{\sum_i v_i x_{io}} \tag{2.1.2}$$

subject to

$$\frac{\sum_r u_r y_{ro}}{\sum_i v_i x_{io}} \leq 1, \quad \forall j = 1, ..., n,$$

$$u_r, v_i \geq 0, \quad \forall i, r.$$

The above ratio form yields an infinite number of solutions. However, the transformation developed by Charnes and Cooper (1962) for linear fractional programming selects a solution (u, v) for which $\sum_{i=1}^{m} \alpha$ = 1 and yields the equivalent linear programming problem in which the change of variables from (u, v) to (μ, ν) is a result of the "Charnes-Cooper" transformation.

$$Max \quad z = \sum_{r=1}^{s} \mu_r y_{ro} \tag{2.1.3}$$

subject to

$$\sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} \nu_i x_{ij} \leq 0,$$

$$\sum_{i=1}^{m} \nu_i x_{io} = 1,$$

$$\mu_r, \nu_i \geq 0,$$

for which the LP dual problem is

$$\theta^* = min \quad \theta \tag{2.1.4}$$

subject to

$$\sum_{j=1}^{n} x_{ij}\lambda_j \leq \theta x_{io}, \quad i = 1, ..., m,$$

$$\sum_{j=1}^{n} y_{ij}\lambda_j \geq y_{ro}, \quad r = 1, ..., s,$$

$$\lambda_j \geq 0, \quad j = 1, ..., n.$$

Furthermore,

$$Max \quad \sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+ \tag{2.1.5}$$

subject to

$$\sum_{r=1}^{s} \lambda_j y_{rj} - s_r^+ = y_{ro}, \quad r = 1, ..., s,$$

$$\sum_{i=1}^{m} \lambda_j x_{ij} - s_i^- = \theta^* x_{io}, \quad i = 1, ..., m,$$

$$\lambda_j, s_i^-, s_r^+ \geq 0, \quad \forall i, j, r,$$

where we note the choices of s_i^- and s_r^+ (slack variables) do not affect the optimal θ^*, which is determined from model (2.1.4). These development lead to the following definition:

Definition 2.1. *(D EA Efficiency). The performance of DMU_o is fully (100%) efficient if and only if both (1) θ^* = 1 and (2) all slacks $s_i^{-*} = s_r^{+*} = 0$.*

Definition 2.2. *(Weakly DEA Efficient) . The performance of DMU_o is weakly efficient if and only if both (1) $\theta^* = 1$ and (2) $s_i^{-*} \neq 0$ and/or $s_r^{+*} \neq 0$ for some i or r in some alternate optima.*

2.2 Review of Mathematical Models

Charnes *et al.* (1978), presented the maximization problem model in a multiple-input multiple-output case by using linear programming. They assume that there are n DMUs to be evaluated. Each DMU consumes m different inputs to produce s different outputs. Specifically, DMU_j

consumes amount x_{ij} of input i and produces amount y_{rj} of output r. They presented the following model:

$$Max \quad h_o(u,v) = \frac{\sum_r u_r y_{ro}}{\sum_i v_i x_{io}} \tag{2.2.1}$$

subject to

$$\frac{\sum_r u_r y_{ro}}{\sum_i v_i x_{io}} \leq 1, \quad \forall j = 1, ..., n,$$

$$u_r, v_i \geq 0, \quad \forall i, r,$$

where y_{ro} = output being observed, x_{io} = input being observed, u_r, v_i are weight associated with output and input respectively.

The above ratio form yields an infinite number of solutions. However, the transformation into linear fractional programming selects a solution λ as the weight, i.e., the solution (u, v) for which $\sum_{i=1}^{m} \lambda = 1$ and yields the equivalent linear programming problem in which the change of variables from (u, v) to (μ, ν) is a result of the transformation,

$$Max \quad \theta = \sum_{r=1}^{s} \mu_r y_{ro} \tag{2.2.2}$$

subject to

$$\sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} \nu_i x_{ij} \leq 0,$$

$$\sum_{i=1}^{m} \nu_i x_{io} = 1,$$

$$\mu_r, \nu_i \geq 0.$$

$$Max \sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+ \tag{2.2.3}$$

subject to

$$\sum_{r=1}^{s} \lambda_j y_{rj} - s_r^+ = y_{ro}, \quad r = 1, ..., s,$$
$$\sum_{i=1}^{m} \lambda_j x_{ij} - s_i^- = \theta^* x_{io}, \quad i = 1, ..., m,$$
$$\lambda_j, s_i^-, s_r^+ \geq 0, \quad \forall i, j, r.$$

These development lead to the following definition: Charnes *et al.* (1978)

Definition 2.3. *(D EA Efficiency). The performance of* DMU_o *is fully (100%) efficient if and only if both (1)* θ^* *= 1 and (2) all slacks* $s_i^{-*} = s_r^{+*} = 0$.

Definition 2.4. *(Weakly DEA Efficient) . The performance of* DMU_o *is weakly efficient if and only if both (1)* $\theta^* = 1$ *and (2)* $s_i^{-*} \neq 0$ *and/or* $s_r^{+*} \neq 0$ *for some i or r in some alternate optima.*

Equation (2.2.3) is called the Charnes, Cooper, and Rhodes (CCR) model, Charnes *et al.* (1978).

Banker *et al.* (1984) added a constraint $\sum_{j=1}^{n} \lambda = 1$ which introduces an extra variable that makes it possible to effect returns-to-scale evaluations (increasing, constant, or decreasing).

Thus, the Banker, Charnes, Cooper (BCC) model is also referred to as the VRS (Variable Returns to Scale) model and distinguished from the CCR model which is referred to as the CRS (Constant Returns to Scale) model.

We now proceed to compare and contrast the input and output orientations of the CCR model.

Fare and Grosskopf (2000) developed a two stage network data envelopment analysis to adress the original DEA that considers the system as a whole(black box). They assume that complex systems exist in real life that have internal structures connected and that the internal structures work independently. They also assume that the internal structures are connected in series and that the output from stage one is an input to stage two, and are known as intermidiate input. We assume each DMU_j (j = 1, 2, . . ., n) has m inputs x_{ij}, (i = 1, 2, . . ., m) to the first stage, and D outputs z_{dj}, (d = 1, 2, . . ., D)from that stage. These D outputs then become the inputs to the second stage and will be referred to as intermediate measures. The outputs from the second stage are y_{rj}, (r = 1, 2, . . ., s). We denote the efficiency for the first stage as e_j^1 and second stage as e_j^2 , for each DMU_j. Using the Constant Returns to Scale (CRS) DEA model of Charnes *et al.* (1978), we define

$$e_j^1 = \frac{\sum_{d=1}^{D} w_d z_{dj}}{\sum_{i=1}^{m} v_i x_{ij}}, \tag{2.2.4}$$

$$e_j^2 = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{d=1}^{D} \hat{w}_d z_{dj}}, \tag{2.2.5}$$

where $v_i, w_d, \hat{w}_d$, and u_r are unknown non-negative weights. Note that w_d can be equal to $\hat{w}_d$.

2.2.1 Efficiency Decomposition Method

It is useful to point out that given individual efficiency measures e_j^1 and e_j^2, for stages 1 and 2 respectively, it is reasonable to define the efficiency of the overall two-stage process either as $\frac{1}{2}(e_j^1 + e_j^2)$ or $e_j^1 * e_j^2$. The above definition ensures that the two-stage process is efficient if and only if $e_j^1 = e_j^2$ = 1.

2.2.2 Other Network DEA

- The examples above, is the case that the intermediate measures are the only inputs to the second stage.
- There are other types of two-stage processes and even DMUs with network structures that may have inputs to the second stage in addition to the intermediate measures.
- The network DEA approach of Fare and Whittaker (1995), Fare and Grosskopf (1996), and the slacks-based network DEA approach of Tone and Tsutsui (2009, 2010) may involve more than two stages. Fukuyama and Weber (2010) considers a slacks-based measure for a two-stage process with bad outputs. Chen et al. (2009) developed a network DEA model incorporating dynamic effects in production networks. A number of empirical studies have used this type of DEA technique.

We should point out that in their paper, the second stage has not only the inputs from the first stage, but also its own inputs not linked with the first stage, i.e. additional inputs to the second stage are introduced. As a results,

$$e_j^2 = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{d=1}^{D} \hat{w}_d z_{dj} + \sum_{h=1}^{H} Q_h x_{hj}^2}, \tag{2.2.6}$$

where x_{hj}^2 (h = 1, ..., H) are inputs to the second stage that are not related to the first stage.

2.2.3 Two Stage Network System with Feed Backs

Liang *et al.* (2008) extended the basic two-stage network system to include situations where outputs from the second stage can be fed back as inputs to the first stage.

- Note that both intermediate variables and the feed-back variables constitute what Cook and Zhu (2006) call dual-role variables, in that they play a role of both output and input simultaneously.

- Liang *et al.* (2008) develop a base model where the (additive) average of two stages efficiency scores is defined as the overall efficiency as in Chen *et al.* (2009). This base average model can be regarded as a centralized model, Liang *et al.* (2008). To adopt the concept of leader-follower approach.

  In addition to the notation used in the previous sections,

we have outputs $f_{gj}(g = 1, 2, ..., p)$ that flow back to stage 1 and become part of the set of inputs to the first stage. We refer to the f_{gj} as feedback variables. The z_{dj} and f_{gj} can then be regarded as dual-role variables. Using additive efficiency decomposition,

$$Max \quad \frac{1}{2}\left(\frac{\sum_{d=1}^{D} h_d z_{do}}{\sum_{i=1}^{m} v_i x_{io} + \sum_{g=1}^{P} w_g f_{go}} + \frac{\sum_{g=1}^{P} w_g f_{go} + \sum_{r=1}^{s} u_r y_{ro}}{\sum_{d=1}^{D} h_d z_{do}}\right) \tag{2.2.7}$$

subject to

$$\frac{\sum_{d=1}^{D} h_d z_{dj}}{\sum_{i=1}^{m} v_i x_{ij} + \sum_{g=1}^{P} w_g f_{gj}} \leq 1,$$

$$\frac{\sum_{g=1}^{P} w_g f_{gj} + \sum_{r=1}^{s} u_r y_{rj}}{\sum_{d=1}^{D} h_d z_{dj}} \leq 1,$$

$$v_i, w_g, h_d, u_r \geq 0,$$

where v_i, u_r, w_g and h_d are unknown weights. Note that it is assumed each stage applies the same weights h_d on the intermediate measures and w_g on the feedback variables f_{gj}. Chung *et al.* (1997) and Fare *et al.* (1994) extended this idea using directional distance functions (DDF) to evaluate efficiency of DMUs. In this approach the desirable outputs can be expanded, and the controllable inputs and undesirable outputs can be reduced based on a given direction vector. In a similar vein, the incorporation of undesirable output can be included through various methodologies including slack-based measures.

- Other approaches describe how to include undesirable output in the context of network DEA Liu et al., (2015), panel data Arabi et al.(2015); Avkiran, (2015); Fare and Grosskopf, (1997); Herrera-Restrepo *et al.*,(2016), and composite indicators Fusco *et al.*, (2020). Li *et al.* (2022) introduced a bargaining game for allocating a fixed cost across decision-making units with undesirable outputs.

- Recently Kao and Hwang (2021, 2023) introduced an innovative approach that disentangles the impact of generating undesirable outputs from the inefficiency associated with producing desirable outputs when assessing the efficiency of a production unit. Literature has reported many examples in DEA considering CO_2 as an undesirable factor, especially with respect to environmental efficiency. This has recently been considered one of the important applications of DEA (Emrouznejad *et al.*, 2023)

2.2.4 Evaluating Environmental efficiency

- Rezaei *et al.* (2022) used network data envelopment analysis to evaluate the economic efficiency and environmental efficiency of European Union Countries.
- In the first stage, the economic efficiency of European Union (EU) countries was calculated according to the input variables.
- In the second stage, the environmental efficiency of EU countries was evaluated based on the two important outputs of greenhouse gas emissions and the absorption of polluting gases. Summarilly, countries were ranked accordingly.

Model

This evaluate the performance of the whole network and its components due to the unfavorable output. The efficiency of each stage can be calculated as follows:

$$E_1 = \frac{\sum\limits_{e=1}^{q} n_{1j} z_{1j}^1}{\sum\limits_{i=1}^{m} \nu_{1j} x_{1j}^1}, \tag{2.2.8}$$

$$E_2 = \frac{\sum\limits_{e=1}^{q} n_{1j} z_{1j}^2}{\sum\limits_{e=1}^{q} n_{1j} z_{1j}^1 + \sum\limits_{r=1}^{s} u_{1j}^b y_{1k}^{b2}}, \tag{2.2.9}$$

$$E_p = \frac{\sum\limits_{e=1}^{q} n_{1j} z_{1j}^p + \sum\limits_{r=1}^{s} u_{rj} y_{1k}^{gp}}{\sum\limits_{e=1}^{q} n_{1j} z_{1j}^1 + \sum\limits_{r=1}^{s} u_{1j}^b y_{1k}^{b2}}. \tag{2.2.10}$$

The performance of the whole network can be written as a convex linear combination as follows:

$$E_{overall} = \sum_{e=1}^{q} w_p E_p, \quad where, \sum_{e=1}^{q} w_p = 1. \tag{2.2.11}$$

Note that the weight of each stage indicates the importance of that stage compared to other network stages. One way to choose the weight for each step is to consider the ratio of the input of each step to the total network inputs, which can be written as follows:

$$W_1 = \frac{\sum\limits_{i=1}^{m} v_{ik}x_{ik}}{\sum\limits_{i=1}^{m} v_{ik}x_{ik} + \sum\limits_{e=1}^{q} n_{ek}z_{ek}^{1} + \sum\limits_{r=1}^{s} u_{rj}^{b}y_{rk}^{b2} + \sum\limits_{r=1}^{s} u_{rj}^{b}y_{rk}^{bp} + \sum\limits_{e=1}^{q} n_{ek}z_{ek}^{p-1}}, \quad (2.2.12)$$

$$W_2 = \frac{\sum\limits_{e=1}^{q} n_{ek}z_{ek}^{1-p} + \sum\limits_{r=1}^{s} u_{rj}^{b}y_{rk}^{b2}}{\sum\limits_{i=1}^{m} v_{ik}x_{ik} + \sum\limits_{e=1}^{q} n_{ek}z_{ek}^{1} + \sum\limits_{r=1}^{s} u_{rj}^{b}y_{rk}^{b2} + \sum\limits_{r=1}^{s} u_{rj}^{b}y_{rk}^{bp} + \sum\limits_{e=1}^{q} n_{ek}z_{ek}^{p-1}}, \quad (2.2.13)$$

$$W_p = \frac{\sum\limits_{e=1}^{q} n_{ek}z_{ek}^{1-p} + \sum\limits_{r=1}^{s} u_{rj}^{b}y_{rk}^{bp}}{\sum\limits_{i=1}^{m} v_{ik}x_{ik} + \sum\limits_{e=1}^{q} n_{ek}z_{ek}^{1} + \sum\limits_{r=1}^{s} u_{rj}^{b}y_{rk}^{b2} + \sum\limits_{r=1}^{s} u_{rj}^{b}y_{rk}^{bp} + \sum\limits_{e=1}^{q} n_{ek}z_{ek}^{p-1}}. \quad (2.2.14)$$

Finally, the performance of the whole network can be written as follows:

$$E_{overall} = \sum_{e=1}^{q} w_p E_p = \frac{\sum\limits_{p=1}^{P}\left(\sum\limits_{e=1}^{q} n_{ek}z_{ek}^{p} + \sum\limits_{r=1}^{s} u_{rj}^{g}y_{rk}^{gp}\right)}{\sum\limits_{i=1}^{m} v_{ik}x_{ik} + \sum\limits_{r=1}^{s} u_{rj}^{b}y_{rk}^{b1} + \sum\limits_{p=2}^{P}\left(\sum\limits_{e=1}^{q} n_{ek}z_{ek}^{p-1} + \sum\limits_{r=1}^{s} u_{rj}^{b}y_{rk}^{bp}\right)}. \quad (2.2.15)$$

$$Max \quad E_{overall} = \sum_{p=1}^{P}\left(\sum_{e=1}^{q} n_{ek}z_{ek}^{p} + \sum_{r=1}^{s} u_{rj}^{g}y_{rk}^{gp}\right) \quad (2.2.16)$$

subject to

$$\sum_{i=1}^{m} v_{ik}x_{ik} + \sum_{r=1}^{s} u_{rj}^{b}y_{rk}^{b1} + \sum_{p=2}^{P}\left(\sum_{e=1}^{q} n_{ek}z_{ek}^{1-p} + \sum_{r=1}^{s} u_{rj}^{b}y_{rk}^{bp}\right) = 1,$$

$$\left(\sum_{r=1}^{s} u_{rj}^{g}y_{rj}^{g1} + \sum_{e=1}^{q} n_{ej}z_{ej}^{1}\right) \leq \sum_{i=1}^{m} v_{ij}x_{ij} + \sum_{r=1}^{s} u_{rj}^{b}y_{rj}^{b},$$

$$\sum_{p=2}^{P}\left(\sum_{r=1}^{s} u_{rj}^{g}y_{rj}^{gp} + \sum_{e=1}^{q} n_{ej}z_{ej}^{p}\right) \leq \sum_{p=2}^{P}\left(\sum_{e=1}^{q} n_{ej}z_{ej}^{1-p} + \sum_{r=1}^{s} u_{rj}^{b}y_{rj}^{bp}\right),$$

$$n_{ej}, u_{rj}, v_{ij} \geq 0.$$

Kidney Allocation Problem

Hamidzadeh *et al.*(2024) presented a novel two-stage network data envelopment analysis model for kidney allocation problem under medical and logistical uncertainty: The research presents a novel method for assessing the efficiency and ranking of qualified organ-patient pairs as decision-making units (DMUs) for kidney allocation problem in the existence of COVID-19 pandemic and uncertain medical and logistical data.

Model

There are n homogeneous decision-making units DMU_j(j = 1, ... , n) which in the first stage I inputs x_{ij}(i = 1, ... , I) and D outputs z_{dj}(d = 1, ... , D) exit. The outputs of the first stage, defined as intermediate products, are considered as the inputs for the second stage, and H inputs f_{hj}(h = 1, ... , H) are also entered into the system in this stage, and ultimately R outputs y_{rj}(r = 1, ... , R) exit the second stage. Moreover, the non-negative weights u_i(i = 1, ... , I), v_d(d = 1, ... , D), w_h(h = 1, ... , H), and π_r (r = 1, ... , R) are delegated to the x_{ij} (i = 1, ... , I) , z_{dj} (d = 1, ... , D) , f_{hj} (h = 1, ... , H) , and y_{rj}(r = 1, ... , R), respectively.

$$Eff_k^{total} = Max \quad \sum_{d=1}^{D} \theta_d \tilde{z}_{dk} + \sum_{r=1}^{R} \tau_r \tilde{y}_{rk} \tag{2.2.17}$$

subject to

$$\sum_{i=1}^{I} \mu_i \tilde{x}_{ik} + \sum_{d=1}^{D} \theta_d \tilde{z}_{dk} + \sum_{h=1}^{H} \omega_h \tilde{f}_{hk} = 1,$$

$$\sum_{d=1}^{D} \theta_d \tilde{z}_{dj} \sum_{i=1}^{I} \mu_i \tilde{x}_{ij} \leq 0 \quad \forall j,$$

$$\sum_{r=1}^{R} \tau_r \tilde{y}_{rj} - \sum_{d=1}^{D} \theta_d \tilde{z}_{dj} - \sum_{h=1}^{H} \omega_h \tilde{f}_{hj} \leq 0, \quad \forall j,$$

$$\mu_i, \theta_d, \omega_h, \tau_r \geq 0, \forall i, r, d, h.$$

In the presented Model (2.2.17), $(\sim)$ demonstrates the epistemic uncertainty in the related parameters. Moreover, the first stage efficiency can be formulated as Model (2.2.18):

$$Eff_k^1 = Max \quad \sum_{d=1}^{D} \theta_d \tilde{z}_{dk} \tag{2.2.18}$$

subject to

$$\sum_{i=1}^{I} \mu_i \tilde{x}_{ik} = 1,$$
$$\sum_{d=1}^{D} \theta_d \tilde{z}_{dj} - \sum_{i=1}^{I} \mu_d \tilde{x}_{ij} \leq 0, \quad \forall j,$$
$$\sum_{r=1}^{R} \tau_r \tilde{y}_{rj} - \sum_{d=1}^{D} \theta_d \tilde{z}_{dj} - \sum_{h=1}^{H} \omega_h f_{hj} \leq 0, \quad \forall j,$$
$$\sum_{d=1}^{D} \theta_d \tilde{z}_{dk} + \sum_{r=1}^{R} \tau_r \tilde{y}_{rk} - Eff_k^{total^*} \sum_{d=1}^{D} \theta_d z_{dk} - Eff_k^{total^*} \sum_{h=1}^{H} \omega_h \tilde{f}_{hk} = Eff_k^{total^*},$$
$$\mu_i, \theta_d, \omega_h, \tau_r \geq 0, \forall i, r, d, h.$$

Subsequently, as shown in Models (2.2.17) and (2.2.18), the objective function and related constraints were modified. However, none of these modifications affect the models optimal solution.

Gerami *et al.* (2023) considered this model when the data is fuzzy.

2.2.5 Uncontrollable Input and Undesirable Output

Taleb *et al.* (2024) build the mathematical formula for the non-oriented DDF and SBM models in the presence of uncontrollable inputs and undesirable outputs, we consider a production system comprising of J DMUs. Each DMU has four factors: controllable inputs, uncontrollable inputs, desirable outputs, and undesirable outputs. The vectors of the controllable and uncontrollable inputs are described as $x^C \in \Re_+^{m_1}, x^{NC} \in \Re_+^{m_2}$, while the vectors of the desirable (good) and undesirable (bad) outputs are described as $y^D \in \Re_+^{s_1}, y^{ND} \in \Re_+^{s_2}$. Because non-oriented efficiency measures have a higher discrimination power in assessing the efficiency of the DMUs, the paper proposes a new mix-efficiency measure based on non-oriented DDF and SBM models with uncontrollable inputs and undesirable outputs. The new DDF and SBM models seek to decrease controllable inputs and undesirable outputs, as well as increase desirable outputs, while preserving uncontrollable inputs at their fixed levels, as defined by the empirical PPS of the DDF and the SBM. In order to consider the technology of DDF with uncontrollable inputs and undesirable outputs, assume $(g_{x^C} \neq 0, g_{x^{NC}} = 0, g_{y^D} \neq 0, g_{y^{ND}} \neq 0)$ yields the directional distance function in the existence of uncontrollable inputs and undesirable outputs. They set the directional vector of uncontrollable inputs to zero because their levels are beyond management control.

$\overrightarrow{DT}(x^C, x^{NC}, y^D, y^{ND}; g) = max[\beta \in [0,1] | (x^C \beta g_{x^C}, x^{NC}, y^D + \beta g_{y^D}, y^{ND} - \beta g_{y^{ND}}) \in T]$.

Taleb *et al.*(2024) presented this paper that, there are two major issues facing the management of decision-making units. The first one is how to deal with uncontrollable inputs whose levels are determined by exogenous fixed factors. The second is how to deal with undesirable outputs that are accompanied by desirable outputs. The effect of the operating environment is frequently captured by uncontrollable inputs and undesirable outputs. The paper presented new directional mix-efficiency measure and slacks-based measure models.

The usefulness and applicability of the proposed models are assessed by measuring the eco-efficiency of the Organization for Economic Co-Operation and Development (OECD) countries. Moreover, they presented the following model:

Parameters

i: 1,..., m_1 index of controllable (i.e., discretionary) inputs.

m_1: number of controllable inputs.

l: 1,..., m_2 index of uncontrollable (i.e., non-discretionary) inputs.

m_2: number of uncontrollable inputs.

r_1: 1,..., s_1 index of desirable (i.e., good) outputs.

s_1: number of desirable outputs.

r_2: 1,..., s_2 index of undesirable (i.e., bad) outputs.

s_2: number of undesirable outputs.

j: 1,..., J index of evaluated DMUs.

o: subscript factor revealing a specific DMU whose efficiency is being measured.

J: number of DMUs whose efficiency is being measured.

x^C_{io}: positive amount of controllable input i of DMUo.

x^{NC}_{lo} : positive amount of uncontrollable input l of DMUo.

y_{r1o}^{G} : positive amount of desirable output r_1 of DMUo.

y_{r2o}^{B} : positive amount of undesirable output r_2 of DMUo.

Variables

$(\eta_1, ..., \eta_J)$: non-negative multipliers used for calculating a reference set of evaluated DMUs in the data set.

a_{io}^{C-} : controllable input slack (i.e., potential reduction) of controllable input i of DMUo.

a_{lo}^{NC} : uncontrollable input slack of uncontrollable input l of DMUo.

$b_{r1o}^{G}+$: desirable output slack (i.e., potential expansion) of desirable output r_1 of DMUo.

b_{r2o}^{B} -: undesirable output slack (i.e., potential reduction) of undesirable output r_2 of DMUo.

non-oriented DDF model Taleb *et al.*(2024) presented A non-oriented DDF model in the presence of uncontrollable inputs and undesirable outputs (DDF-NCIUO) for evaluating DMUo as formulated below

$$min \quad \tau_o^{NCIUO} = \frac{1-\beta_o}{1+\beta_o} \tag{2.2.19}$$

subject to

$$\sum_{j=1}^{J} x_{ij}^{C}\eta_j \leq x_{io}^{C}(1-\beta_o), \quad i = 1, ..., m_1,$$

$$\sum_{j=1}^{J} x_{ij}^{NC}\eta_j \leq x_{io}^{NC}, \quad i = 1, ..., M_2,$$

$$\sum_{j=1}^{J} y_{r_1j}^{G}\eta_j \leq y_{r_1o}^{G}(1+\beta_o), \quad r_1 = 1, ..., S_1,$$

$$\sum_{j=1}^{J} y_{r_2j}^{B}\eta_j \leq y_{r_2o}^{B}(1-\beta_o), \quad r_2 = 1, ..., S_2,$$

$$\sum_{j=1}^{J} \eta_j = 1,$$

$$\eta_j \geq 0, \quad j = 1, ..., J,$$

$$\beta_o \in [0, 1),$$

$$x_{io}^{C}, x_{lo}^{NC}, y_{r_1o}^{G}, y_{r_2o}^{B} \in \Re^{m_1+m_2+s_1+s_2}.$$

where β_o is the proportional rate of decrease in input and increase in output.

2.2.6 Non-oriented SBM Model

Taleb *et al.*(2024) modified the SBM model of Tone (2003) by incorporating uncontrollable inputs into a specific input constraint of the model. The slack of uncontrollable inputs is omitted from the models target function because the efficiency evaluation only depends on controllable variables, whereas the slack of uncontrollable inputs can be considered in its relevant constraint to avoid the infeasibility issue. Thus, under the VRS technology, an SBM for the case of uncontrollable inputs and undesirable outputs (SBM-NCIUO) is proposed for evaluating DMUo, as follows:

$$\delta_0^{NCIUO} = min \frac{1 - \frac{1}{m_1}\left(\sum_{i=1}^{m_1} \frac{a_i^{C-}}{x_{io}^{C}}\right)}{1 + \frac{1}{s_1+s_2}\left(\sum_{r_1=1}^{s_1} \frac{b_{r_1}^{G+}}{y_{r_1 0}^{G}} + \sum_{r_2=1}^{s_2} \frac{b_{r_2}^{B-}}{y_{r_2 0}^{B}}\right)} \tag{2.2.20}$$

subject to

$$\sum_{j=1}^{J} x_{ij}^{C}\eta_j = x_{io}^{C} - a_i^{C}, \quad i = 1, ..., m_1,$$

$$\sum_{j=1}^{J} x_{ij}^{NC}\eta_j = x_{io}^{NC} - a_i^{NC-}, \quad i = 1, ..., M_2,$$

$$\sum_{j=1}^{J} y_{r_1 j}^{G}\eta_j = y_{r_1 0}^{G+} + b_{r_1}^{G+}, \quad r_1 = 1, ..., S_1,$$

$$\sum_{j=1}^{J} y_{r_2 j}^{B}\eta_j = y_{r_2 o}^{B} - b_{r_2}^{B-}, \quad r_2 = 1, ..., S_2,$$

$$a_{i0}^{C-} \geq 0, a_i^{NC-} \geq 0, b_{r_1}^{G+} \geq 0, b_{r_2}^{B-} \geq 0,$$

$$\sum_{j=1}^{J} \eta_j = 1,$$

$$\eta_j \geq 0, \quad j = 1, ..., J,$$

$$x_{io}^{C}, x_{lo}^{NC}, y_{r_1 o}^{G}, y_{r_2 0}^{B} \in \Re^{m_1+m_2+s_1+s_2}.$$

The computed ratio of efficiency scores resulting from model (2.2.19) and model (2.2.20) efficiency scores represents a new directional mix-efficiency measure with uncontrollable inputs and undesirable outputs (MIX-NCIUO). MIX-NCIUO of DMUo (Ψ_o^{NCIUO}) is calculated as

$$\Psi_0^{NCIUO} = \frac{\delta_0^{NCIUO}}{\tau_0^{NCIUO}}, \qquad \tau_0^{NCIUO} \neq 0. \tag{2.2.21}$$

CHAPTER THREE

Methodology

3.1 Introduction

Data envelopment analysis (DEA) is categorized into two main approaches: black-box and network DEA.
Black-box DEA utilizes initial inputs to generate final outputs without looking into internal stages, thereby overlooking inefficiencies.
Conversely, the network DEA considers internal stages to identify the sources of inefficiency for each DMU.

The introduction of a network structure within DEA is to address the limitations of the black-box DEA model. This enhanced model, known as network DEA (NDEA), incorporates network structures to capture more refine efficiency considerations. Specifically, when NDEA comprises precisely two stage network systems or two internal processes, it is termed a two-stage DEA model. The network structure can manifest in various forms, including series, parallel and mix structures. In a two-stage DEA model, the first stage employs inputs to generate intermediate measures, which then serve as inputs for the second stage to produce final outputs.

3.2 Directional Distance Function (DDF)

Notation

Consider n decision-making units (DMUs), indexed by $j = 1, \ldots, n$, each represented as s-stage network. For the focal DMU_j, we observe:

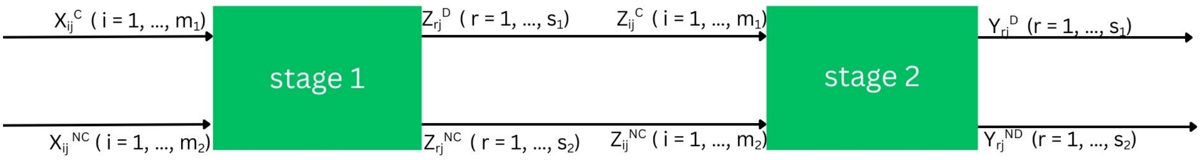

Figure 3.1: Two stage network system

- Stage $s \in \{1, ..., E\}$ inputs: $\boldsymbol{x}_j^{(s)} \in \mathbb{R}_+^{m_s}$, partitioned into

$$\boldsymbol{x}_j^{(s)} = \big(\boldsymbol{x}_j^{(s),C},\, \boldsymbol{x}_j^{(s),U}\big),$$

  where C marks controllable inputs and U marks uncontrollable inputs.

- Stage s desirable outputs: $\boldsymbol{y}_j^{(s)} \in \mathbb{R}_+^{r_s}$.

- Stage s undesirable outputs: $\boldsymbol{b}_j^{(s)} \in \mathbb{R}_+^{q_s}$.

- Intermediates linking the stages: $\boldsymbol{z}_j^{(s)} \in \mathbb{R}_+^{\ell}$ produced by Stage 1 and used by Stage 2 and up to stage E.

Directions policy: For each stage s, fix direction vectors

$$\boldsymbol{g}_{x_i}^{(s)} = \big(\boldsymbol{g}_{x_i}^{(s),C}, \boldsymbol{g}_{x_i}^{(s),U}\big), \quad \boldsymbol{g}_{y_r}^{(s)}, \quad \boldsymbol{g}_{b_k}^{(s)}, \quad \boldsymbol{g}_{z_l},$$

with $\boldsymbol{g}_{y_r}^{(s)}, \boldsymbol{g}_{z_l} \geq 0$ (expansion directions), and $\boldsymbol{g}_{x_i}^{(s)}, \boldsymbol{g}_{b_k}^{(s)} \geq 0$ (contraction directions). For uncontrollable inputs, set the corresponding direction components to zero.

Network Technology (Convex Representation)

DEA assumes the production technology (the feasible set of input-output combinations) can be approximated by combinations of observed DMUs. Formally, if DMU A and B are feasible, then any weighted average (convex combination) of A and B is also feasible. This is the convexity assumption.

(x,y) $\in$ T, $(x^{'}, y^{'})$ $\in$ T $\implies$ λ(x, y) + (1-λ)($x^{'}, y^{'}$) $\in$ T, $\forall\lambda \in$ [0, 1].

So the objective function (e.g., maximize expansion/contraction of desirable outputs/inputs) is always subject to feasibility defined via λ.

Relation to the Objective Function

1. The objective function pushes the DMU towards the efficient frontier by maximizing β (or minimizing slack sum in SBM).

2. The convexity constraint ($\sum_j \lambda_j$ = 1, $\lambda_j \geq 0$) ensures that the frontier being targeted is exactly the convex hull of real DMUs (under VRS).

3. The free disposability inequalities make sure the adjustment (β) is compared against something feasible (you can always waste inputs/bad output or drop good outputs).

4. Thus, the convex combination conditions on λ anchor the objective function to a feasible benchmark frontier.

3.2.1 Directional Distance Function (DDF) on a Network

For DMU_j and a chosen direction mix, the network DDF seeks the largest scalar β such that the point

$$\left(\boldsymbol{x}_j^{(s)} - \beta \boldsymbol{g}_{x_i}^{(s)},\ \boldsymbol{y}_j^{(s)} + \beta \boldsymbol{g}_{y_r}^{(s)},\ \boldsymbol{b}_j^{(s)} - \beta \boldsymbol{g}_{b_k}^{(s)}\right)_{s=1,\dots,E}, \quad \boldsymbol{z}_j + \beta \boldsymbol{g}_{z_l}$$

belongs to the feasible network technology and respects the linking constraints. The value β measures the maximal simultaneous contraction of controllable inputs/undesirable output and expansion of desirables/intermediates outputs along the chosen directions.

Given the directional distance function for a technology set T, β is defined as :

f (x, y, b, z: g) = max{β: (x - βg_{x_i}, y + βg_{y_r}, b-βg_{b_k}, z - βg_{z_l}) $\in$ T}

where,

x $\in \Re_+^m$ = input vector,

y $\in \Re_+^r$ = desirable output vector, b $\in \Re^q_+$ = undesirable output vector,

g = (g_x, g_y, g_b, g_z) is the directional vector, specifically the direction along which inputs/ undesirable output are reduced and desirable output are expanded.

T is the production possibility set.

so, β is the maximum feasible expansion/ contraction factor along the chosen direction vector g.

f is the largest step-size along g that keeps the adjusted point feasible.

Primal LP for the s-Stage Network DDF

Solve the following LP for each DMU_j, where j = 1, ..., n. Let $\mathbb{S}$ be the set of all admissible strategy for the primal LP.

Model (VRS)

at stage s

where s = 1, ..., E

$$\max_{\boldsymbol{\lambda}^{(s)} \in \mathbb{S}} \quad \beta^{(s)}(\lambda^{(s)}) \tag{3.2.1}$$

$$\text{subject to} \tag{3.2.2}$$

$$\sum_{j=1}^{n} \lambda_j^{(s)} \, x_{j,i}^{(s)} \;\leq\; x_{i,o}^{(s)} - \beta \, g_{x_i}^{(s)}, \qquad \forall i \in \mathcal{I}_s^C \cup \mathcal{I}_s^U, \tag{3.2.3}$$

$$\sum_{j=1}^{n} \lambda_j^{(s)} \, y_{j,r}^{(s)} \;\geq\; y_{r,o}^{(s)} + \beta \, g_{y_r}^{(s)}, \qquad \forall r \in \mathcal{R}_s, \tag{3.2.4}$$

$$\sum_{j=1}^{n} \lambda_j^{(s)} \, b_{j,k}^{(s)} \;\leq\; b_{k,o}^{(s)} - \beta \, g_{b_k}^{(s)}, \qquad \forall k \in \mathcal{Q}_s, \tag{3.2.5}$$

$$\sum_{j=1}^{n} \lambda_j^{(s)} \, z_{j,\ell} \;=\; z_{\ell,o} + \beta \, g_{z_l}^{(s)}, \qquad \forall \ell \in \mathcal{L}, \tag{3.2.6}$$

$$\sum_{j=1}^{n} \lambda_j^{(s)} = 1, \qquad \lambda_j^{(s)} \geq 0. \tag{3.2.7}$$

Discussion.

Constraints (3.2.3)–(3.2.5) enforce feasibility of the projected point for each stage under the chosen directions. The intermediate link is enforced by (3.2.6); if one wishes to allow loss or storage between stages, replace equality by $\geq$ or $\leq$ as appropriate. VRS is imposed by (3.2.7).

Uncontrollable Inputs Uncontrollable inputs cannot be freely reduced. The standard treatments is:

Zero-direction. For all $i \in \mathcal{I}_s^U$, set $g_{x_i}^{(s)} = 0$. This preserves the observed level of uncontrollable while allowing movement on controllable.

Undesirable Outputs

Undesirables are modelled for *reduction* along $\boldsymbol{g}_{b_k}^{(s)}$ through constraints (3.2.5).

Interpretation

A larger $\beta^{(s)}$ indicates greater potential to *simultaneously*

(i) reduce controllable inputs and undesirable outputs and

(ii) expand desirable outputs and intermediates.

3.2.2 Non-Oriented DDF model

:

A non-oriented DDF model in the presence of uncontrollable inputs and undesirable outputs (DDF-NCIUO) for evaluating DMU_j at stage s, (s = 1, ..., E)

$$\tau_s^{NCIUO} = \frac{1-\beta^{(s)}}{1+\beta^{(s)}}. \tag{3.2.8}$$

Proposition 4. DMU_j is Pareto-Koopmans efficient in the DDF-NCIUO model if and only if its directional input and output vector values are zero (i.e., $\beta^{(s)} = 0$). This condition is equivalent to the efficiency score of DMU_j being equal to one (i.e.,$\tau_s^{NCIUO} = 1$). If the directional vector value lies in the interval (0, 1), then DMU_j is inefficient.

3.3 Non-oriented Slack Based Measure Model Incorporating Uncontrollable Inputs and Undesirable Outputs

Variable Definitions

Consider a set of n decision making units (DMUs), indexed by $j = 1, \ldots, n$. Each DMU consists of E connected sub-processes (stages) s, (s = 1, ..., E).

- $x_{ij}^{c(s)}$: amount of controllable input i used by DMU j in Stage s, $i = 1, \ldots, m_s^c$.

- $x_{ij}^{u(s)}$: amount of uncontrollable input i used by DMU j in Stage s, $i = 1, \ldots, m_s^u$.
- $y_{rj}^{(s)}$: desirable output r produced by Stage s, $r = 1, \ldots, v_s$.
- $b_{kj}^{(s)}$: undesirable output k produced by Stage s, $k = 1, \ldots, q_s$.
- z_{pj} : intermediate product p linking Stages, $p = 1, \ldots, l$.

Slack variables:

- $s_i^{-c(s)}$: slack for controllable input i in Stage s.
- $s_i^{-u(s)}$: slack for uncontrollable input i in Stage s (fixed at zero).
- $s_r^{+y(s)}$: slack for desirable output r in Stage s.
- $s_k^{-b(s)}$: slack for undesirable output k in Stage s.
- $s_p^{-z(s)}$: slack for intermediate variable l in stage s
- Intensity variables: $\lambda_j^{(s)}$ for j=1,…,n

Stage s technology set:

$$T^{(s)} = \left\{ \left(x_j^{c(s)}, x_j^{u(s)}\right) \to \left(y_j^{(s)}, b_j^{(s)}, z_j\right) \;\middle|\; \begin{array}{l} \sum_{j=1}^{n} \lambda_j x_{ij}^{c(s)} + s_i^{-c(s)} = x_{io}^{c(s)}, \quad i = 1, \ldots, m_s^c, \\ \sum_{j=1}^{n} \lambda_j x_{ij}^{u(s)} = x_{io}^{u(s)}, \quad i = 1, \ldots, m_s^u, \\ \sum_{j=1}^{n} \lambda_j y_{rj}^{(s)} - s_r^{+y(s)} = y_{ro}^{(s)}, \quad r = 1, \ldots, v_s, \\ \sum_{j=1}^{n} \lambda_j b_{rj}^{(s)} + s_r^{-b(s)} = b_{ro}^{(s)}, \quad r = 1, \ldots, q_s, \\ \sum_{j=1}^{n} \lambda_j z_{lj} + s_l^{-z} = z_{lo}^{(s)}, \quad l = 1, \ldots, l, \\ \lambda_j \geq 0 \end{array} \right\}$$

where,

$$x_j^{c(s)} \in \mathbb{R}_+^{m_{sc}}, \quad x_j^{u(s)} \in \mathbb{R}_+^{m_{su}}, \quad y_j^{(s)} \in \mathbb{R}_+^{r_1}, \quad b_j^{(s)} \in \mathbb{R}_+^{q_1},$$

Stage s SBM for DMU$_j$

$$\rho_s^{(NCIUO)} = \frac{1 - \frac{1}{m_{sc}}\left(\sum_{i=1}^{m_{sc}} \frac{s_i^{-c(s)}}{x_{i,o}^{c(s)}}\right)}{1 + \frac{1}{r_s + q_s}\left(\sum_{i=1}^{r_s} \frac{s_r^{+y(s)}}{y_{r,o}^{(s)}} \; + \; \sum_{r=1}^{q_s} \frac{s_r^{-b(s)}}{b_{r,o}^{(s)}}\right)}. \tag{3.3.1}$$

subject to (3.3.2)

$$\sum_{j=1}^{n} \lambda_j x_{ij}^{c(s)} + s_i^{-c(s)} = x_{io}^{c(s)}, \quad i = 1, \ldots, m_s^c, \tag{3.3.3}$$

$$\sum_{j=1}^{n} \lambda_j x_{ij}^{u(s)} = x_{io}^{u(s)}, \quad i = 1, \ldots, m_s^u, \tag{3.3.4}$$

$$\sum_{j=1}^{n} \lambda_j y_{rj}^{(s)} - s_r^{+y(s)} = y_{ro}^{(s)}, \quad r = 1, \ldots, v_s, \tag{3.3.5}$$

$$\sum_{j=1}^{n} \lambda_j b_{rj}^{(s)} + s_r^{-b(s)} = b_{ro}^{(s)}, \quad r = 1, \ldots, q_s, \tag{3.3.6}$$

$$\sum_{j=1}^{n} \lambda_j z_{lj} + s_l^{-z} = z_{lo}^{(s)}, \quad l = 1, \ldots, l, \tag{3.3.7}$$

$$\lambda_j \geq 0. \tag{3.3.8}$$

3.3.1 Linearization of the stage-s SBM by the Charnes–Cooper transformation

Consider the s stage SBM for DMU$_o$:

$$\rho_s^{(NCIUO)} = \frac{1 - \dfrac{1}{m_s^c}\left(\sum_{i=1}^{m_s^c} \frac{s_i^{-c(s)}}{x_{i,o}^{c(s)}}\right)}{1 + \dfrac{1}{r_s + q_s}\left(\sum_{r=1}^{r_s} \dfrac{s_r^{+y(s)}}{y_{r,o}^{(s)}} + \sum_{r=1}^{q_s} \dfrac{s_r^{-b(s)}}{b_{r,o}^{(s)}}\right)} \tag{3.3.9}$$

subject to (3.3.10)

$$\sum_{j=1}^{n} \lambda_j\, x_{ij}^{c(s)} + s_i^{-c(s)} = x_{i,o}^{c(s)}, \qquad i = 1, \ldots, m_s^c, \tag{3.3.11}$$

$$\sum_{j=1}^{n} \lambda_j\, x_{ij}^{u(s)} = x_{i,o}^{u(s)}, \qquad i = 1, \ldots, m_s^u, \tag{3.3.12}$$

$$\sum_{j=1}^{n} \lambda_j\, y_{rj}^{(s)} - s_r^{+y(s)} = y_{r,o}^{(s)}, \qquad r = 1, \ldots, r_s, \tag{3.3.13}$$

$$\sum_{j=1}^{n} \lambda_j\, b_{rj}^{(s)} + s_r^{-b(s)} = b_{r,o}^{(s)}, \qquad r = 1, \ldots, q_s, \tag{3.3.14}$$

$$\sum_{j=1}^{n} \lambda_j\, z_{\ell j} + s_\ell^{-z} = z_{\ell,o}^{(s)}, \qquad \ell = 1, \ldots, L, \tag{3.3.15}$$

$$\sum_{j=1}^{n} \lambda_j = 1, \tag{3.3.16}$$

$$\lambda_j \geq 0, \quad s_i^{-c(s)},\ s_r^{+y(s)},\ s_r^{-b(s)},\ s_\ell^{-z} \geq 0. \tag{3.3.17}$$

Charnes–Cooper transformation.

Define,

$$D \;=\; 1 + \frac{1}{r_s + q_s}\left(\sum_{r=1}^{r_s} \frac{s_r^{+y(s)}}{y_{r,o}^{(s)}} + \sum_{r=1}^{q_s} \frac{s_r^{-b(s)}}{b_{r,o}^{(s)}}\right), \qquad t \;=\; \frac{1}{D}.$$

Introduce scaled variables

$$\tilde{\lambda}_j = \lambda_j\, t, \qquad \tilde{s}_i^{-c} = s_i^{-c(s)}\, t, \qquad \tilde{s}_r^{+y} = s_r^{+y(s)}\, t, \qquad \tilde{s}_r^{-b} = s_r^{-b(s)}\, t, \qquad \tilde{s}_\ell^{-z} = s_\ell^{-z}\, t.$$

Because $D\,t = 1$ we obtain the linear normalization:

$$t + \frac{1}{r_s + q_s}\left(\sum_{r=1}^{r_s} \frac{\tilde{s}_r^{+y}}{y_{r,o}^{(s)}} + \sum_{r=1}^{q_s} \frac{\tilde{s}_r^{-b}}{b_{r,o}^{(s)}}\right) = 1. \tag{3.3.18}$$

Multiply constraints (3.3.11)–(3.3.15) by t and substitute scaled variables to obtain linear equalities:

$$\sum_{j=1}^{n} \tilde{\lambda}_j\, x_{ij}^{c(s)} + \tilde{s}_i^{-c} = t\, x_{i,o}^{c(s)}, \qquad i = 1, \ldots, m_s^c, \tag{3.3.19}$$

$$\sum_{j=1}^{n} \tilde{\lambda}_j\, x_{ij}^{u(s)} = t\, x_{i,o}^{u(s)}, \qquad i = 1, \ldots, m_s^u, \tag{3.3.20}$$

$$\sum_{j=1}^{n} \tilde{\lambda}_j\, y_{rj}^{(s)} - \tilde{s}_r^{+y} = t\, y_{r,o}^{(s)}, \qquad r = 1, \ldots, r_s, \tag{3.3.21}$$

$$\sum_{j=1}^{n} \tilde{\lambda}_j\, b_{rj}^{(s)} + \tilde{s}_r^{-b} = t\, b_{r,o}^{(s)}, \qquad r = 1, \ldots, q_s, \tag{3.3.22}$$

$$\sum_{j=1}^{n} \tilde{\lambda}_j\, z_{\ell j} + \tilde{s}_\ell^{-z} = t\, z_{\ell,o}^{(s)}, \qquad \ell = 1, \ldots, L. \tag{3.3.23}$$

$$\sum_{j=1}^{n} \tilde{\lambda}_j = t. \tag{3.3.24}$$

Linearized LP. The original objective $\rho_s^{(NCIUO)} = N/D$ transforms into a linear objective in the scaled variables:

$$\rho_s^{(NCIUO)} = N \cdot t = t - \frac{1}{m_s^c} \sum_{i=1}^{m_s^c} \frac{\tilde{s}_i^{-c}}{x_{i,o}^{c(s)}}.$$

Hence the Charnes–Cooper linear program for stage s (DMU$_o$) is:

Maximize

$$\Phi \;=\; t \;-\; \frac{1}{m_s^c}\sum_{i=1}^{m_s^c}\frac{\tilde{s}_i^{-c}}{x_{i,o}^{c(s)}} \tag{3.3.25}$$

subject to : (3.3.26)

$$\sum_{j=1}^{n}\tilde{\lambda}_j\, x_{ij}^{c(s)} + \tilde{s}_i^{-c} = t\, x_{i,o}^{c(s)}, \qquad i = 1,\dots,m_s^c, \tag{3.3.27}$$

$$\sum_{j=1}^{n}\tilde{\lambda}_j\, x_{ij}^{u(s)} = t\, x_{i,o}^{u(s)}, \qquad i = 1,\dots,m_s^u, \tag{3.3.28}$$

$$\sum_{j=1}^{n}\tilde{\lambda}_j\, y_{rj}^{(s)} - \tilde{s}_r^{+y} = t\, y_{r,o}^{(s)}, \qquad r = 1,\dots,r_s, \tag{3.3.29}$$

$$\sum_{j=1}^{n}\tilde{\lambda}_j\, b_{rj}^{(s)} + \tilde{s}_r^{-b} = t\, b_{r,o}^{(s)}, \qquad r = 1,\dots,q_s, \tag{3.3.30}$$

$$\sum_{j=1}^{n}\tilde{\lambda}_j\, z_{\ell j} + \tilde{s}_\ell^{-z} = t\, z_{\ell,o}^{(s)}, \qquad \ell = 1,\dots,L, \tag{3.3.31}$$

$$t + \frac{1}{r_s + q_s}\left(\sum_{r=1}^{r_s}\frac{\tilde{s}_r^{+y}}{y_{r,o}^{(s)}} + \sum_{r=1}^{q_s}\frac{\tilde{s}_r^{-b}}{b_{r,o}^{(s)}}\right) = 1, \tag{3.3.32}$$

$$\sum_{j=1}^{n}\tilde{\lambda}_j = t, \tag{3.3.33}$$

$$\tilde{\lambda}_j \ge 0, \quad \tilde{s}_i^{-c},\ \tilde{s}_r^{+y},\ \tilde{s}_r^{-b},\ \tilde{s}_\ell^{-z} \ge 0, \quad t \ge 0. \tag{3.3.34}$$

Recovering original variables. If $(\tilde{\lambda}^*, \tilde{s}^*, t^*)$ is an optimal solution with $t^* > 0$, recover

$$\lambda_j^* = \frac{\tilde{\lambda}_j^*}{t^*}, \qquad s_i^{-c(s)*} = \frac{\tilde{s}_i^{-c*}}{t^*}, \qquad s_r^{+y(s)*} = \frac{\tilde{s}_r^{+y*}}{t^*}, \qquad s_r^{-b(s)*} = \frac{\tilde{s}_r^{-b*}}{t^*},$$

and the SBM score is

$$\rho_s^{(NCIUO)*} = \Phi^* = t^* - \frac{1}{m_s^c}\sum_{i=1}^{m_s^c}\frac{\tilde{s}_i^{-c*}}{x_{i,o}^{c(s)}}.$$

Remark. Care must be taken with zero denominators $x_{i,o}^{c(s)}, y_{r,o}^{(s)}, b_{r,o}^{(s)}$; handle zeros by appropriate conventions (variable selection, small positive perturbation, or alternative normalization) and document the choice.

3.4 Network Mix Efficiency

The network mix efficiency scores are calculated as :

$$\Psi_s^{NCIUO} = \frac{\rho_s^{NCIUO}}{\tau_s^{NCIUO}}, \qquad \tau_s^{NCIUO} \neq 0, \qquad (3.4.1)$$
$$\forall \quad s = 1, ..., E.$$

Overall efficiency score

$$\theta^{NCIUO} = \sum_{s=1}^{E} \lambda_s \Psi_s^{NCIUO}, \qquad (3.4.2)$$

where λ is the weight, $\sum_{s=1}^{E} \lambda = 1$.

CHAPTER FOUR

Results and Discussion

4.1 Descriptive Results

To see what makes our model different from the black box (BB) model, we present the data set of 5 OEDC countries as follows:

Table 4.1: Data set of the 5 OEDC countries.

DMU	X^{C_1}	X^{C_2}	X^{C_3}	X^U	Z^D	Z^{UD}	Z^C	Y^D	Y^{UD}
Australia	111.12	156.53	97.6	534	252232	571839.85	824071.9	850.32	381.36
Canada	179.46	63.62	228.3	537	626000	724625.05	1350625.1	1424.06	560.8
France	286.2	22.43	197.9	867	5524606	510830.35	6035436.4	2582.39	375.68
Germany	415.9	281.44	241.6	700	6216656	942526.35	7159182.4	3323.81	787.24
Japan	666.9	207.58	503.7	1668	1043845	1324225.51	2368071	4356.33	1251.17

source: column 1-5, 9-10 (Taleb *et al.*, 2024)

where,

DMU = Country, X^{C_1} = Labour force(10^5), X^{C_2} = Coal consumption$(10^6 tons)$, X^{C_3} = Petroleum consumption$(10^4 barrels/day)$, X^U = Precipitation(millimetre /year), Z^D = Power generated(KW), $Z^{UD} = CO_2$ during power generation $(10^6 tons)$, Z^C = intermediate value, Y^D = GDP $(10^9 US\$)$, $Y^{UD} = CO_2 (10^6 tons)$ after mitigation.

Table 4.2: Two-Stage DEA Efficiency Results for Five DMUs

DMU	BB	N	S1	S2	Stage-1 Gap	Stage-2 Gap	Bottleneck
Australia	0.3735	0.3503	0.0892	0.6114	91.1%	38.9%	Stage 1 major
Canada	0.4425	0.5007	0.1475	0.8538	85.2%	14.6%	Stage 1 major
France	1.0000	0.9259	1.0000	0.8517	0.0%	14.8%	Stage 2 only
Germany	1.0000	0.9507	0.9392	0.9622	6.1%	3.8%	Mild S1 & S2.
Japan	1.0000	0.6689	0.3377	1.0000	66.2%	0.0%	Stage 1 only

4.2 Interpretation of Results

Table 4.2 presents the efficiency results for the five DMUs using both the black-box (BB) approach and the two-stage network DEA model. The results reveal several important insights:

1. **DMU1 and DMU2:** These units exhibit extremely low Stage 1 efficiency scores (0.0892 and 0.1475, respectively), indicating that the primary source of inefficiency lies in the first stage. Stage 2 performance is comparatively better, but overall performance remains constrained by the first stage.

1. **DMU3:** While Stage 1 efficiency is perfect (1.0000), Stage 2 efficiency (0.8517) is the main source of inefficiency. Improvements should therefore target the second stage.
2. **DMU4:** This DMU is nearly efficient in both stages (0.9392 and 0.9622), with only minor improvements needed to reach the efficiency frontier.
3. **DMU5:** Although appearing fully efficient in the black-box model (1.0000), the two-stage decomposition reveals severe inefficiency in Stage 1 (0.3377) while Stage 2 is fully efficient (1.0000). This highlights hidden inefficiency that is masked in the black-box approach.

Overall, the two-stage network model provides a more detailed picture of efficiency performance, enabling managers to pinpoint the exact stage where improvements are required rather than relying solely on aggregated black-box results.

Why a two-stage network efficiency can exceed the black-box efficiency

- The black-box LP forces one intensity vector to satisfy both input compression and final-output replication. This is often restrictive.
- The two-stage network LP(s) allow stage-specific intensity vectors; this extra degree of freedom enlarges the feasible set for stage-wise problems and may lead to strictly better (smaller) input-contraction factors at the stage level.

Thus the observation for DMU2, namely

$$BB_2 = 0.4425 < N_2 = 0.5007,$$

is consistent with a network formulation that allows stage-specific peer combinations or weighting schemes. Practically, this indicates DMU2 can be represented more favorably (relative to peers) when internal structure is recognized, because one stage provides very favorable peer comparisons that the single-stage (black-box) model cannot jointly exploit.

4.2.1 Data and variables

Data presentation

To see the application and usefulness of the proposed efficiency models, we consider the agricultural production and its environmental impact in Africa. The data sets on African countries for the year 2022 were extracted from the following sources:

1. World Food and Agricultural Organization (https://www.fao.org/faostat/en/home).
2. United Nations Statistical Year book 2024 edition.
3. https://theswiftest.com/biodiversity-index.
4. Individual countries statistical year book.

The status of Decision Making Unit (DMU) was assigned to each African country and its efficiency measures were determined by a two stage network model where agricultural practices was considered at the first stage and its environmental impact was considered at the second stage. The classification of factors are as follows:

stage one

1. X^C controllable inputs to include:
   (a) Labour (Thousands)
   (b) Fertilizer, pesticides, insecticides (Thousand tones)
   (c) land used for irrigation (Thousand HA)
2. X^{NC} uncontrollable inputs which is climate change
   (a) rainfall (volumes)
   (b) temperature (degree Celsius)

   Z^D desirable output to include:
   (a) Income from agricultural products to include crop yield, Timber and fisheries (USD)

   Z^{ND} undesirable outputs to include:
   (a) Carbon emissions (million tones)

stage two

1. Z^C controllable inputs to include:
   (a) output from stage one
   (b) resources allocated for environmental impact (from donor agencies).
   (c) organic manure from agricultural products.
2. Z^{NC} uncontrollable inputs which is natural disaster to include:
   (a) flood (amount of damage)
   (b) droughts and storm

   Y^D desirable output to include:
   (a) biodiversity (thousand)
   (b) GDP (USD)

   Y^{ND} undesirable outputs to include:

(a) loss of biodiversity,

(b) number of mal-nourished,

Table 4.3: Data set of the 54 African countries.

DMU	X^c	X^u	Z^d	Z^{ud}	Z^C	Z^{uc}	Y^d	Y^{ud}
Algeria	2654.2	81.72	3523.9	73	3739.9	511	199.49	57.36
Angola	7259.7	1079.33	1186	119.7	1280	246	242.86	168.97
Benin	1530.1	1367.09	320	25.7	1226	712	136.72	86.48
Botswana	178.7	51.61	303	65.5	393	1	37	73.37
Burkina Faso	6030.3	970.26	859.9	60.4	2360.9	217	63.43	60.54
Burundi	4719.6	1103.47	138.2	13.9	711.2	99	254.63	71.43
Cabo Verde	26.8	172.57	19.2	67.4	102.2	38.439	133.39	21.5
Cameroon	4836.4	3025.66	982	0.4	2170	47	1268.58	174.01
Central African Republic	1353.9	1435.66	308.4	73.6	991.4	86.3	134.57	108.45
Chad	3886.4	395.2	3848.4	207.1	4548.4	18	85	71.47
Comoros	75	3151	35.9	0.4	174.9	1	187.14	23.39
Congo	620.8	2485	150.7	20.7	853.7	39	224.02	129.81
Congo D R	18648.1	127	696.3	693.4	4422.3	227	764.07	250.9
Cote d'Ivoire	4897.3	1135	526	50.8	2571	12.9	351.28	101.95
Djibouti	22.1	320.2	27.8	1.8	86.8	73	130.4	39.62
Egypt	11388.7	3625	9921.6	121.6	15854.6	72	220.54	77.38
Equatorial Guinea	291.7	2463.16	126.9	4.6	137.9	1	299.66	75.08
Eritrea	1076.3	375.94	168	10.5	230	1	178.49	62.71
Eswatini	109.2	835.17	72.6	3.7	178.6	2.058	55.96	49.33
Ethiopia	37720.1	946.26	5367.5	310	58842	61	275.95	155.51
Gabon	211.2	3015	106.9	8.8	239.9	8	616.41	104.52
Gambia	432.3	1126	133	3.1	337	12	147.57	80.85
Ghana	5985.9	1290.15	1013.1	32.7	2133.1	15	314.12	116.66
Guinea	2556.3	3757	732.6	61.4	1233.6	139	451.87	154.88
Guinea0Bissau	372.7	2026	112.5	6.1	260.5	26	157.61	65.75
Kenya	7997.8	959.31	5113.2	120.4	7929.2	21	719.63	198.42
Lesotho	247.6	734	189.8	2.7	341.87	1	27.29	27.68
Liberia	934.6	1799.8	88.5	15.6	575.5	56.28	368.13	87
Libya	602.8	348	408	15.7	672	1	96	41.27
Madagascar	11667.5	1172.5	816.3	66.1	1932.3	199	3770.62	173.8

Table 4.4: Data set of the 54 African countries continue

DMU	X^c	X^u	Z^d	Z^{ud}	Z^C	Z^{uc}	Y^d	Y^{ud}
Malawi	4947.3	1023	909.7	28.6	2279.7	133	181.61	99.32
Mali	5683.4	417.04	606.2	78.3	1824.2	551	110.98	91.12
Mauritania	386.2	128.9	1069.8	22.7	1396.8	39	151.4	67.15
Mauritius	58.3	503.5	107.6	1.9	400.6	1	295.96	33.05
Morocco	5733.3	279.72	4954.3	66.3	6388.3	1	289.65	69.86
Mozambique	10079.7	359	809.5	104.3	3398.5	831	574.88	161.63
Namibia	188.1	721	632.4	27.2	816.4	8	103.62	111.71
Niger	7208.17	391	984.5	80.4	2791.5	228	83.18	58.62
Nigeria	27466.6	337	23432.8	308.1	27024.8	677	485.3	175.25
Rwanda	2473	33.6	318.9	13.2	1653.9	53	222.02	84.38
Sao Tome and Principe	21.1	8746.5	8.3	0.3	89.3	16	154.38	17.24
Senegal	1261.3	30.36	1133.3	36.4	2525.3	1397	199.43	90.91
Seychelles	0.2	2111	143.4	0.1	143.4	1	463.61	23.1
Sierra Leone	1221	4427	246.5	12.2	1151.5	9	355.2	27.4
Somalia	862.5	357.91	680.4	65.5	3087.4	79	292.26	113.49
South Africa	6149.3	511	8225	137.3	9312	65.5	908.41	212.84
South Sudan	2298.3	926.7	3099.8	122.3	5218.8	227	71.03	50
Sudan	6488	47803.5	4151.4	191.7	8043.4	135	192.55	32.9
Tanzania	19832.3	1034.05	4871.4	221.5	7802.4	6	1619.64	228.7
Togo	955	1116.9	118.2	10.9	628.2	186.96	149.99	125.31
Tunisia	1148.1	159.63	1930.9	23.3	3265.9	41	139.9	87.75
Uganda	11767.8	310.13	2676.3	93.9	5745.3	40	340.29	154.05
Zambia	4147.2	1055.61	996.6	116.7	2093.6	51	140.49	120.84
Zimbabwe	3334.5	927.23	1406.3	37.4	2393.3	1	144.29	100.32

The number of evaluated DMUs in a DEA model should, according to the common rule of thumb, be at least three times larger than the total number of inputs and outputs to allow the model differentiate between efficient and inefficient DMUs based on their input-output mixes and to identify the best performers. Otherwise, the problem of discrimination may arise.

More so, DEA models are based on the assumption that the relationship between the input and output is linear (Taleb et al. (2019)).

Table (4.5) below shows the correlation coefficient between the inputs and the outputs.

Table 4.5: Correlation coefficient between inputs and outputs

	X^c	X^u	Z^d	Z^{ud}	Z^C	Z^{uc}	Y^d	Y^{ud}
Xc	1.000	-0.004	0.625	0.724	0.853	0.169	0.319	0.600
Xu	-0.004	1.000	0.059	0.089	0.043	-0.054	-0.031	-0.194
Zd	0.625	0.059	1.000	0.442	0.602	0.217	0.099	0.336
Zud	0.724	0.089	0.442	1.000	0.497	0.156	0.168	0.577
ZC	0.853	0.043	0.602	0.497	1.000	0.085	0.047	0.314
Zuc	0.169	-0.054	0.217	0.156	0.085	1.000	0.007	0.153
Yd	0.319	-0.031	0.099	0.168	0.047	0.007	1.000	0.519
Yud	0.600	-0.194	0.336	0.577	0.314	0.153	0.519	1.000

1. Interpretation

1. Correlation coefficient ranges from -1 to +1
2. +1 indicates Perfect positive relationship: both variables move in the same direction
3. -1 indicates Perfect negative relationship: one increases, the other decreases
4. 0 indicates No linear relationship
5. Values above 0.7 = strong positive correlation
6. Values 0.3 to 0.7 = moderate
7. Below 0.3 = weak

2 Variable Pair Correlation Interpretation

1. X^c & Z^c = 0.85 Strong Stage 1 controllable input (X^c) is strongly related to Stage 2 controllable input (Z^c). DMUs with high inputs in Stage 1 tend to have high in Stage 2 too.

2. X^c & Z^{ud} = 0.72 Strong More inputs (X^c) in Stage 1 implying More undesirable outputs Z^{ud}.

3. X^c & Z^d = 0.62 Moderate More inputs in Stage 1 indicates More desirable outputs Z^d, but less strongly than undesirables.

4. X^c & Y^{ud} = 0.60 Moderate Stage 1 inputs relate to undesirable outputs in Stage 2.

5. Z^d & Z^{ud} = 0.44 Moderate Desirable and undesirable outputs in Stage 1 move together.

6. Y^d & Y^{ud} = 0.52 Moderate In Stage 2, desirable & undesirable outputs rise together.

7. X^u (uncontrollable) correlations Mostly low (less than 0.1) Uncontrollable inputs Xu have weak relationships with most variables - DMUs with higher Xu do not necessarily produce more or less outputs.

3 Implications for Efficiency Models

1. X^c, Z^c, Z^{ud}, Y^{ud} are highly correlated indicates these will drive efficiency scores the most.

2. X^u weakly correlated implying may have less effect unless constraints force them in.

3. Outputs correlation (Z^d, Y^d with Y^{ud}) means high production also brings pollution or undesirable outputs.

Table 4.6: Summary Statistics of Variables

Variable	Min	Max	Mean	StdDev
X^c	0.20	37720.10	4852.679630	7156.970882
X^u	30.36	47803.50	2131.210000	6501.232678
Z^d	8.30	23432.80	1866.846296	3663.198933
Z^{ud}	0.10	693.40	75.135185	112.480898
Z^C	86.80	58842.00	4000.077778	8812.319627
Z^{uc}	1.00	1397.00	142.693278	257.885983
Y^d	27.29	3770.62	362.620370	557.283804
Y^{ud}	17.24	250.90	97.167222	56.461730

The characteristics of the data set for the 54 African countries is presented in 4.6.

Interpretation for Efficiency

1. Large StdDev / wide Min-Max gap implies DMUs differ significantly in resource usage & production.

2. Inputs with very high dispersion (e.g., X^c, Z^c) may dominate efficiency scores unless normalized or weighted properly.

3. Undesirable outputs being small scale means their impact could be minor unless explicitly weighted in the model.

Table 4.7 shows the result of the beta values, efficiencies from stage 1 and stage 2.

Table 4.7: DDF two stage efficiency measures of the 54 African countries.

DMU	$\beta^{(1)}$	τ_1^{NCIUO}	$\beta^{(2)}$	τ_2^{NCIUO}
Algeria	0.000	1.000	0.642	0.218
Angola	0.789	0.118	0.865	0.073
Benin	0.772	0.129	0.781	0.123
Botswana	0.000	1.000	0.636	0.222
Burkina Faso	0.737	0.151	0.715	0.166
Burundi	0.866	0.072	0.686	0.186
Cabo Verde	0.000	1.000	0.128	0.773
Cameroon	0.000	1.000	0.148	0.742
Central African Republic	0.861	0.074	0.825	0.096
Chad	0.181	0.694	0.759	0.137
Comoros	0.750	0.143	0.012	0.976
Congo	0.876	0.066	0.830	0.093
Congo D R	0.962	0.019	0.749	0.143
Cote d\Ívoire	0.805	0.108	0.712	0.168
Djibouti	0.000	1.000	0.000	1.000
Egypt	0.000	1.000	0.722	0.161
Equatorial Guinea	0.762	0.135	0.000	1.000
Eritrea	0.601	0.249	0.384	0.445
Eswatini	0.580	0.265	0.228	0.629
Ethiopia	0.640	0.220	0.839	0.088
Gabon	0.864	0.073	0.000	1.000
Gambia	0.446	0.383	0.686	0.186
Ghana	0.556	0.286	0.767	0.132
Guinea	0.731	0.155	0.742	0.148
Guinea0Bissau	0.848	0.083	0.592	0.256
Kenya	0.254	0.595	0.528	0.309
Lesotho	0.082	0.849	0.165	0.716
Liberia	0.947	0.027	0.624	0.231
Libya	0.452	0.377	0.440	0.389
Madagascar	0.792	0.116	0.000	1.000

Table 4.8: DDF two stage efficiency measures of the 54 African countries continue.

DMU	$\beta^{(1)}$	τ_1^{NCIUO}	$\beta^{(2)}$	τ_2^{NCIUO}
Malawi	0.525	0.311	0.794	0.115
Mali	0.831	0.092	0.801	0.110
Mauritania	0.000	1.000	0.714	0.167
Mauritius	0.000	1.000	0.301	0.537
Morocco	0.025	0.951	0.601	0.250
Mozambique	0.832	0.092	0.711	0.169
Namibia	0.000	1.000	0.818	0.100
Niger	0.751	0.142	0.706	0.172
Nigeria	0.000	1.000	0.766	0.133
Rwanda	0.000	1.000	0.743	0.147
Sao Tome and Principe	0.667	0.200	0.000	1.000
Senegal	0.000	1.000	0.769	0.131
Seychelles	0.000	1.000	0.000	1.000
Sierra Leone	0.724	0.160	0.186	0.686
Somalia	0.558	0.284	0.774	0.127
South Africa	0.000	1.000	0.667	0.200
South Sudan	0.058	0.891	0.655	0.208
Sudan	0.368	0.462	0.409	0.420
Tanzania	0.564	0.278	0.000	1.000
Togo	0.807	0.107	0.825	0.096
Tunisia	0.000	1.000	0.783	0.122
Uganda	0.473	0.358	0.805	0.108
Zambia	0.765	0.133	0.841	0.086
Zimbabwe	0.408	0.421	0.770	0.130

Table 4.7 reveals that Algeria, Botswana, Cabo Verd, Cameroon, Djibouti, Egypt, Seychelles and south Africa and few others are fully efficient in the first stage and Djibouti, Equatorial Guinea, Gabon, Seychelles and Tanzania are fully efficient at the second stage.

Table 4.9 is generated by solving the SBM model in s stage. No country is fully efficient in both stages and this is a confirmation that unlike the DDF model that is bias, SBM is not bias as it considers the slack of both inputs and outputs. Only a hand full of countries are close to the efficiency frontier. Therefore, more attention is to be given to the countries that are inefficient to make them efficient.

Table 4.9: SBM two stage efficiency measures of the 54 African countries

DMU	$\rho_1^{(NCIUO)}$	ranking 1	$\rho_2^{(NCIUO)}$	ranking 2
Algeria	0.632	7	0.080	31
Angola	0.098	38	0.054	39
Benin	0.102	36	0.045	45
Botswana	0.148	29	0.035	52
Burkina Faso	0.118	33	0.029	54
Burundi	0.049	50	0.127	18
Cabo Verde	0.016	53	0.260	11
Cameroon	0.495	9	0.267	10
Central African Republic	0.054	49	0.046	44
Chad	0.288	17	0.037	50
Comoros	0.069	47	0.380	5
Congo	0.063	48	0.071	35
Congo D R	0.018	52	0.101	25
Cote dÍvoire	0.084	43	0.121	20
Djibouti	0.153	26	0.185	14
Egypt	0.645	6	0.077	32
Equatorial Guinea	0.126	31	0.360	8
Eritrea	0.121	32	0.199	13
Eswatini	0.146	30	0.064	37
Ethiopia	0.152	27	0.049	43
Gabon	0.084	42	0.363	7
Gambia	0.178	24	0.087	29
Ghana	0.200	22	0.097	27
Guinea	0.106	35	0.111	23
Guinea0Bissau	0.100	37	0.112	22
Kenya	0.413	11	0.124	19
Lesotho	0.361	14	0.045	46
Liberia	0.041	51	0.169	15
Libya	0.268	18	0.114	21
Madagascar	0.089	41	0.767	2

Table 4.10: SBM two stage efficiency measures of the 54 African countries continue.

DMU	$\rho_1^{(NCIUO)}$	ranking 1	$\rho_2^{(NCIUO)}$	ranking 2
Malawi	0.211	20	0.058	38
Mali	0.075	44	0.034	53
Mauritania	0.689	4	0.075	33
Mauritius	0.399	12	0.427	3
Morocco	0.704	3	0.224	12
Mozambique	0.072	45	0.102	24
Namibia	0.375	13	0.041	47
Niger	0.113	34	0.038	49
Nigeria	0.765	2	0.067	36
Rwanda	0.192	23	0.088	28
Sao Tome and Principe	0.008	54	0.376	6
Senegal	0.423	10	0.051	40
Seychelles	0.561	8	0.985	1
Sierra Leone	0.093	40	0.409	4
Somalia	0.164	25	0.082	30
South Africa	0.683	5	0.135	17
South Sudan	0.352	15	0.035	51
Sudan	0.151	28	0.138	16
Tanzania	0.204	21	0.320	9
Togo	0.072	46	0.050	42
Tunisia	0.884	1	0.050	41
Uganda	0.247	19	0.073	34
Zambia	0.098	39	0.040	48
Zimbabwe	0.307	16	0.099	26

Table 4.11: overall efficiency measure of the 54 African countries.

DMU	Ψ_1^{NCIUO}	Ψ_2^{NCIUO}	θ^{NCIUO}
Algeria	0.631850152	0.367489351	0.499669752
Angola	0.828331822	0.743723398	0.78602761
Benin	0.793919365	0.367769321	0.580844343
Botswana	0.147845755	0.155803389	0.151824572
Burkina Faso	0.781151149	0.174347647	0.477749398
Burundi	0.685539795	0.68224855	0.683894172
Cabo Verde	0.015624906	0.336751945	0.176188426
Cameroon	0.495012082	0.359967948	0.427490015
Central African Republic	0.71968314	0.480294358	0.599988749
Chad	0.415329603	0.268021733	0.341675668
Comoros	0.47995484	0.38950414	0.43472949
Congo	0.952900329	0.762725408	0.857812869
Congo D R	0.924763666	0.707383588	0.816073627
Cote d'Ivoire	0.777171428	0.720971313	0.74907137
Djibouti	0.152745114	0.185176479	0.168960796
Egypt	0.644552021	0.475002168	0.559777094
Equatorial Guinea	0.934096317	0.359927031	0.647011674
Eritrea	0.484076055	0.447127245	0.46560165
Eswatini	0.551364172	0.101360778	0.326362475
Ethiopia	0.690118976	0.554208701	0.622163838
Gabon	0.86932	0.362857206	0.616089
Gambia	0.46551583	0.468054633	0.466785231
Ghana	0.701801217	0.737749953	0.719775585
Guinea	0.68206591	0.747325478	0.714695694
Guinea0Bissau	0.821314	0.437055873	0.629185
Kenya	0.694889735	0.403280361	0.549085048
Lesotho	0.425592458	0.062817757	0.244205107
Liberia	0.67167	0.731855052	0.701765
Libya	0.711919689	0.29403319	0.502976439
Madagascar	0.762534543	0.766982198	0.76475837

Table 4.12: overall efficiency measure of the 54 African countries continue.

DMU	Ψ_1^{NCIUO}	Ψ_2^{NCIUO}	θ^{NCIUO}
Malawi	0.678768917	0.502537301	0.590653109
Mali	0.819042738	0.311522221	0.56528248
Mauritania	0.689306144	0.447606613	0.568456379
Mauritius	0.399427498	0.794560126	0.596993812
Morocco	0.739566235	0.898491256	0.819028745
Mozambique	0.781394139	0.600917996	0.691156067
Namibia	0.374909914	0.409317857	0.392113882
Niger	0.796656102	0.220700816	0.508678459
Nigeria	0.764621042	0.508081201	0.636351121
Rwanda	0.192169201	0.595692929	0.393931065
Sao Tome and Principe	0.037641191	0.376325183	0.206983187
Senegal	0.422851759	0.389998786	0.406425273
Seychelles	0.560654367	0.98473788	0.772696124
Sierra Leone	0.578795039	0.596734953	0.587764996
Somalia	0.576749879	0.646969425	0.611859652
South Africa	0.682870503	0.673432686	0.678151595
South Sudan	0.39484688	0.167057321	0.2809521
Sudan	0.3279296067	0.327847103	0.327888354
Tanzania	0.733004577	0.32005187	0.526528224
Togo	0.670538305	0.518792706	0.594665505
Tunisia	0.884081256	0.411979034	0.648030145
Uganda	0.690138597	0.6734057067	0.681772152
Zambia	0.734573047	0.46258392	0.598578483
Zimbabwe	0.730328369	0.760475417	0.745401893

Table 4.13: overall efficiency measure ranking of the 54 African countries.

DMU	Ψ_1^{NCIUO}	Ψ_2^{NCIUO}	θ^{NCIUO}	Rank
Congo	0.952900329	0.762725408	0.857812869	1
Morocco	0.739566235	0.898491256	0.819028745	2
Congo D R	0.924763666	0.707383588	0.816073627	3
Angola	0.828331822	0.743723398	0.78602761	4
Seychelles	0.560654367	0.98473788	0.772696124	5
Madagascar	0.762534543	0.766982198	0.76475837	6
Cote d’Ivoire	0.777171428	0.720971313	0.74907137	7
Zimbabwe	0.730328369	0.760475417	0.745401893	8
Ghana	0.701801217	0.737749953	0.719775585	9
Guinea	0.68206591	0.747325478	0.714695694	10
Liberia	0.67167	0.731855052	0.701765	11
Mozambique	0.781394139	0.600917996	0.691156067	12
Burundi	0.685539795	0.68224855	0.683894172	13
Uganda	0.690138597	0.673405706	0.681772152	14
South Africa	0.682870503	0.673432686	0.678151595	15
Tunisia	0.884081256	0.411979034	0.648030145	16
Equatorial Guinea	0.934096317	0.359927031	0.647011674	17
Nigeria	0.764621042	0.508081201	0.636351121	18
Guinea0Bissau	0.821314	0.437055873	0.629185	19
Ethiopia	0.690118976	0.554208701	0.622163838	20
Gabon	0.86932	0.362857206	0.616089	21
Somalia	0.576749879	0.646969425	0.611859652	22
Central African Republic	0.71968314	0.480294358	0.599988749	23
Zambia	0.734573047	0.46258392	0.598578483	24
Mauritius	0.399427498	0.794560126	0.596993812	25
Togo	0.670538305	0.518792706	0.594665505	26
Malawi	0.678768917	0.502537301	0.590653109	27

Table 4.14: Ranking of the 54 African countries continue.

DMU	Ψ_1^{NCIUO}	Ψ_2^{NCIUO}	θ^{NCIUO}	rank
Sierra Leone	0.578795039	0.596734953	0.587764996	28
Benin	0.793919365	0.367769321	0.5808443437	29
Mauritania	0.689306144	0.447606613	0.568456379	30
Mali	0.819042738	0.311522221	0.56528248	31
Egypt	0.644552021	0.475002168	0.559777094	32
Kenya	0.694889735	0.403280361	0.549085048	33
Tanzania	0.733004577	0.32005187	0.526528224	34
Niger	0.796656102	0.220700816	0.508678459	35
Libya	0.711919689	0.29403319	0.502976439	36
Algeria	0.631850152	0.367489351	0.499669752	37
Burkina Faso	0.781151149	0.174347647	0.477749398	38
Gambia	0.46551583	0.468054633	0.466785231	39
Eritrea	0.484076055	0.447127245	0.46560165	40
Comoros	0.47995484	0.38950414	0.43472949	41
Cameroon	0.495012082	0.359967948	0.427490015	42
Senegal	0.422851759	0.389998786	0.406425273	43
Rwanda	0.192169201	0.595692929	0.393931065	44
Namibia	0.374909914	0.40931785	0.392113882	45
Chad	0.415329603	0.268021733	0.341675668	46
Sudan	0.327929606	0.327847103	0.3278883547	47
Eswatini	0.551364172	0.101360778	0.326362475	48
South Sudan	0.39484688	0.167057321	0.2809521	49
Lesotho	0.425592458	0.062817757	0.244205107	50
Sao Tome and Principe	0.037641191	0.376325183	0.206983187	51
Cabo Verde	0.015624906	0.336751945	0.176188426	52
Djibouti	0.152745114	0.185176479	0.168960796	53
Botswana	0.147845755	0.155803389	0.151824572	54

Final Observations

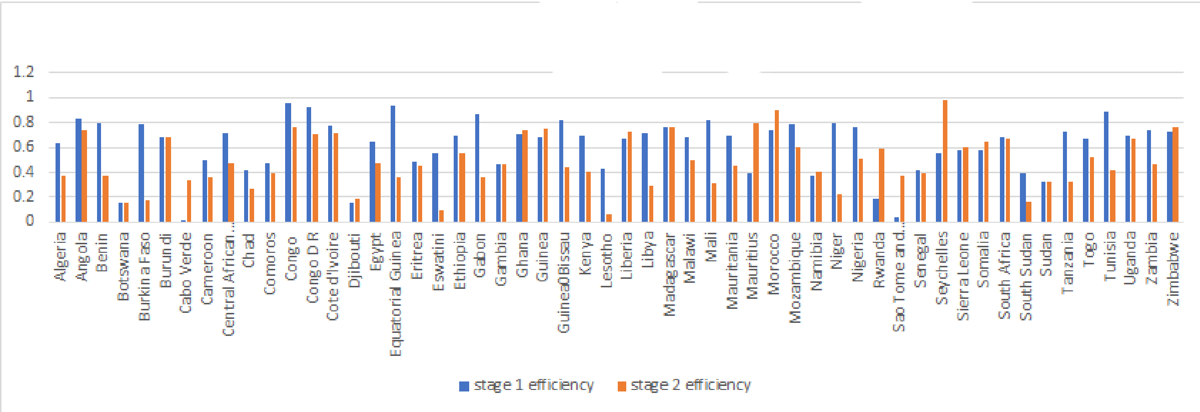

Figure 4.1: Bar chart for stage 1 and 2 efficiency measure

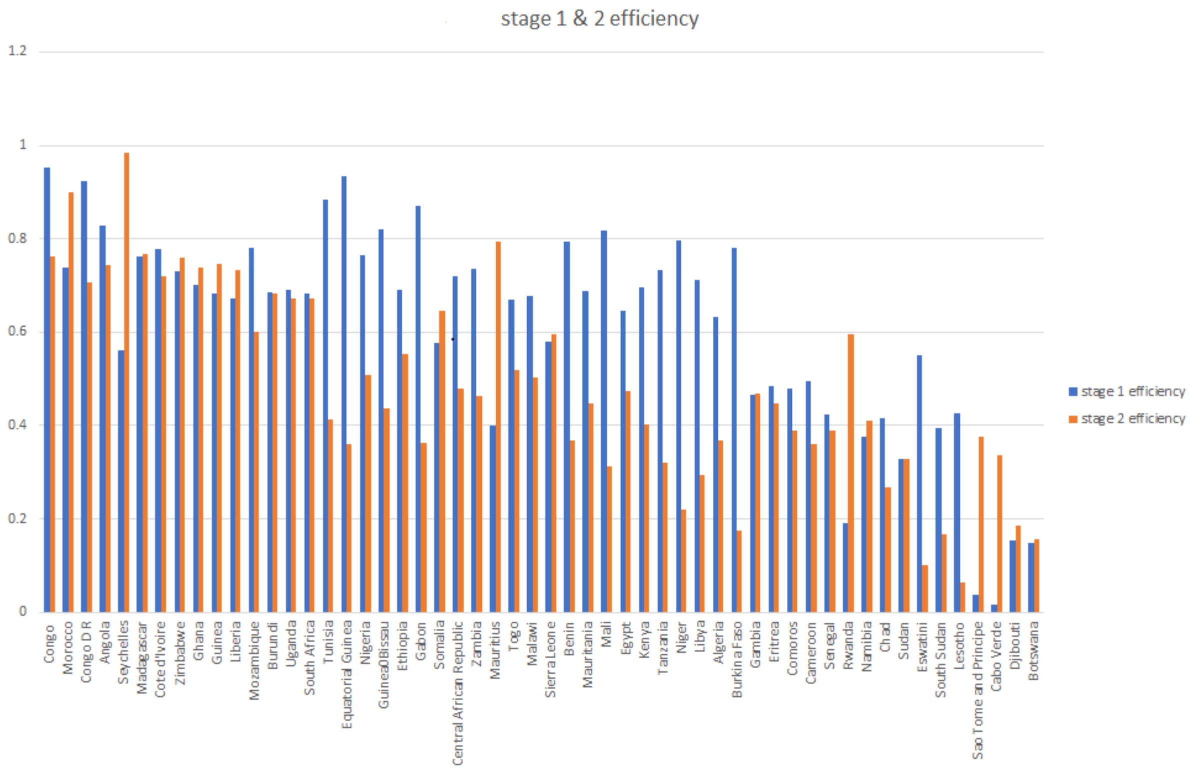

Figure 4.2: Bar chart for stage 1 and 2 ranking

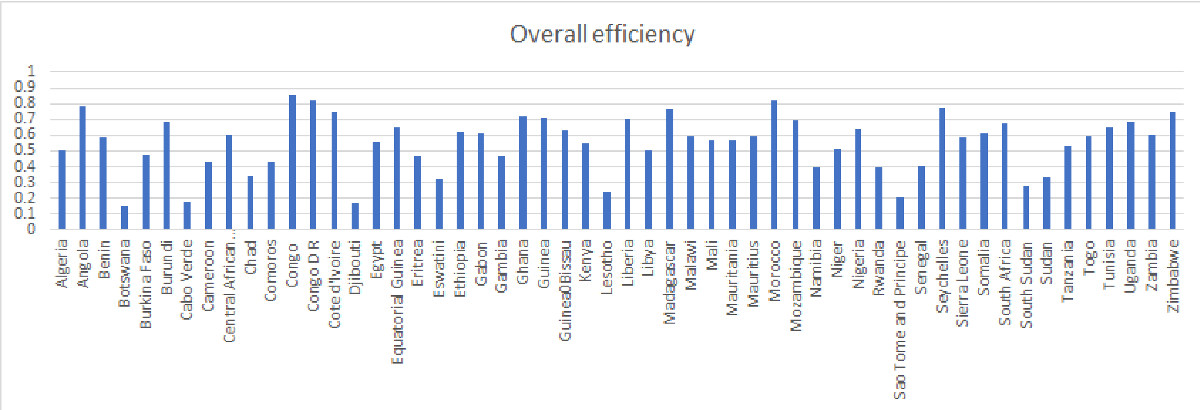

Figure 4.3: Bar chart for overall efficiency measure

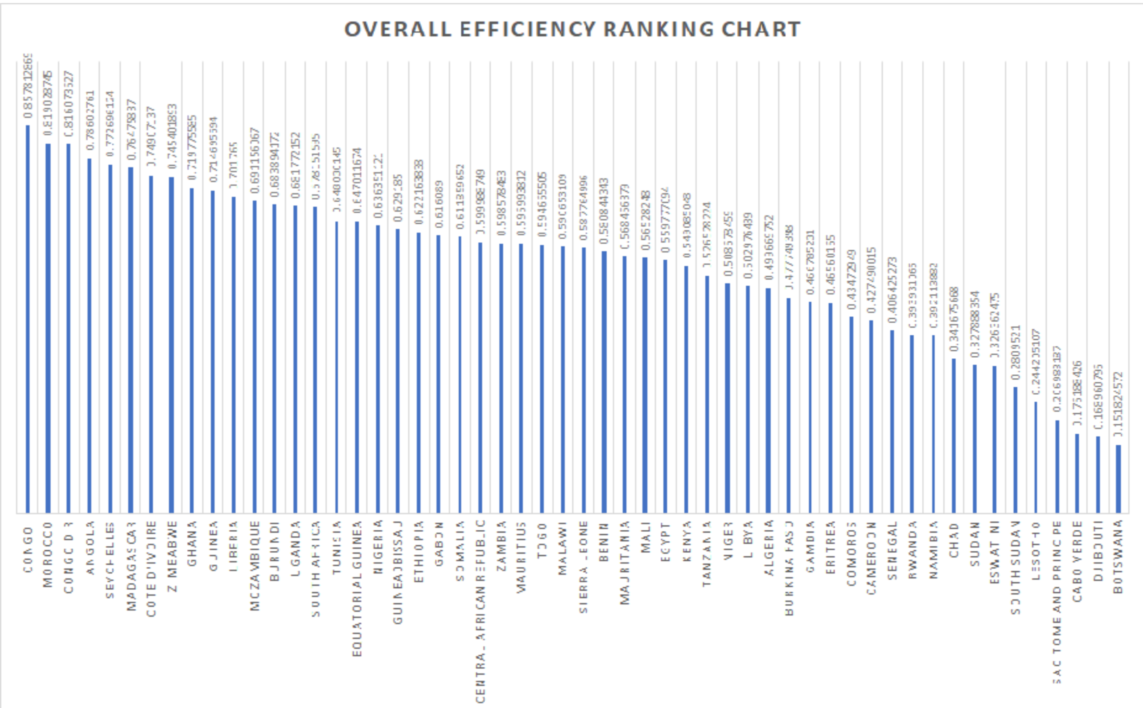

Figure 4.4: Bar chart for overall ranking

Top Efficiency Leaders: no country is seen with full efficiency.

→ however, some countries like Congo, Morocco, Congo D R, Angola and few others are Consistently strong across both stages.

Stage-imbalanced countries: Equatorial guinea, Tunisia, Nigeria (strong Stage 1, weak Stage 2).

Lagging countries: Botswana, Djibouti, cabo verde and others.

→ Weak across both stages.

Overall pattern:

- Only a handful of countries achieve near-perfect efficiency.
- Most countries show a *Stage gap* (either strong at foundations but weak in transformation, or vice versa).
- The overall ranking rewards countries with **balanced development** across both stages.

CHAPTER FIVE

Summary and Conclusion

5.1 Summary

This thesis examines efficiency measurement in complex production systems by applying non-radial network Data Envelopment Analysis (DEA) models to evaluate the performance of African countries in the context of agricultural sustainability. Conventional DEA models treat decision-making units (DMUs) as black boxes and ignore their internal production structures, which may lead to misleading efficiency estimates, particularly in the presence of uncontrollable inputs and undesirable outputs. In the agricultural sector, such simplifications overlook the sequential relationship between production activities and their environmental consequences.

To address these limitations, this study develops and applies a two-stage network DEA framework based on the Directional Distance Function (DDF) and the Slacks-Based Measure (SBM). In the proposed structure, the first stage represents agricultural production, where conventional agricultural inputs are transformed into desirable agricultural outputs. The second stage captures the environmental impact of agricultural activities, where outputs from the first stage act as intermediate products and give rise to undesirable environmental outputs. The two stages are linked through intermediate variables, ensuring consistency of material and technological flows across the system.

The network DDF model measures inefficiency by allowing the expansion of desirable agricultural outputs and the contraction of undesirable environmental outputs along a specified directional vector, while restricting uncontrollable inputs from adjustment. The network SBM model evaluates efficiency using normalized slack variables, providing a non-radial assessment of inefficiency at both the stage and system levels. In both frameworks, undesirable outputs are incorporated explicitly without transformation, preserving the integrity of the production technology.

Using data from African countries, the proposed models generate efficiency scores for agricultural production efficiency, environmental efficiency, and overall system efficiency. Based on the overall efficiency measures, African countries are ranked according to their ability to achieve agricultural output while minimizing environmental damage. The empirical results reveal substantial heterogeneity in performance across countries and demonstrate that high agricultural production efficiency does not necessarily imply strong environmental performance.

To the best of our knowledge, such ranking have not been considered in the literature on network mix directional efficiency measure in the presence of controllable and uncontrollable inputs as well as desirable and undesirable outputs. Therefore, we can say that controllable and uncontrollable inputs as well as desirable and undesirable outputs have been integrated into a standard network DDF and SBM models.

5.2 Findings

The findings from this study are as follows:

- The network model provides a more detailed result of efficiency performance, enabling managers to pinpoint the exact stage where improvements are required rather than relying solely on aggregated black-box results.
- Both stage wise efficiency and the overall efficiency measures consider controllable and uncontrollable inputs as well as desirable and undesirable outputs.

5.3 Contribution to knowledge

The contribution to knowledge includes:

- A system that computes the efficiencies of internal stages and the whole system.
- A comprehensive quantitative approach that measure system efficiency by using both controllable/ uncontrollable inputs and desirable/ undesirable outputs.

5.4 Conclusion

The first research question, which concerns the incorporation of uncontrollable inputs and undesirable outputs into non-radial network DDF and SBM models, has been addressed through theoretically consistent model formulations. In the first stage, uncontrollable inputs associated with natural or structural conditions are treated as exogenous factors and are excluded from efficiency improvement directions. In the second stage, environmental pollutants generated by agricultural activities are modeled explicitly as undesirable outputs, whose reduction contributes to efficiency improvement. These formulations ensure fairness, feasibility, and consistency in efficiency measurement.

The second research question, which focuses on the evaluation of efficiencies at both the component and system levels, has been resolved by deriving stage-specific efficiency measures alongside an overall system efficiency index. The linking constraints between stages ensure that inefficiencies in agricultural production are transmitted to environmental performance, thereby capturing the trade-offs inherent in sustainable agricultural development. This dual-level evaluation provides richer diagnostic information than conventional DEA models and supports more targeted policy interventions.

The empirical application to African countries shows that efficiency rankings differ significantly when internal structures and environmental impacts are explicitly considered. Some countries perform efficiently in agricultural production but exhibit poor environmental efficiency, while others achieve a more balanced performance across both stages. These findings highlight the importance of integrated efficiency assessment for designing coherent agricultural and environmental policies.

In conclusion, this thesis advances DEA methodology by integrating two-stage network structures, non-radial DDF and SBM models, uncontrollable inputs, and undesirable outputs into a unified analytical framework. The application to African agriculture demonstrates the practical relevance of the proposed approach and provides valuable insights for evaluating sustainable agricultural performance and informing evidence-based policy decisions.

5.5 Recommendation for further studies

Despite its contributions, this study has certain limitations. The proposed models are deterministic and assume precise data, which may not fully capture uncertainty in environmental measurement. Future research may extend the framework to dynamic, stochastic, or fuzzy network DEA models, and incorporate time-series data to assess changes in sustainability performance over time. Further empirical applications using alternative environmental indicators or regional analyses within Africa would also enhance the robustness of the findings.

References

Arabi, B., Munisamy, S., and Emrouznejad, A. (2015). A new slacks-based measure of Malmquist-Luenberger index in the presence of undesirable outputs. *Omega*, 51: 29–37.

Avkiran, N. K. (2015). An illustration of dynamic network DEA in commercial banking including robustness tests. *Omega*, 55: 141–150.

Banker, R. D., Charnes, A., and Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, 30(9): 1078–1092.

Centre for Research on the Epidemiology of Disasters (CRED) (2022). Institute Health and Society – UCLouvain B.1.30.15, Clos Chapelle aux Champs 1200 Brussels, Belgium. www.emdat.be

Charnes, A., Cooper, W. W., and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6): 429–444.

Charnes, A. and Cooper, W. (1962). Programming with linear fractional functionals. *Naval Research Logistics Quarterly*, 9: 181–186.

Chen, W., Wang, Y. and Yang, S. (2009). Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 199–208).

Chen, Y., Cook, W. D., Li, N., and Zhu, J. (2009). Additive efficiency decomposition in two-stage DEA. *European Journal of Operational Research*, 196: 1170–1176.

Cook, W. D., and Zhu, J. (2006). Rank order data in DEA: A general framework. *European Journal of Operational Research*, 174(2): 1021–1038.

Chung, Y. H., Fare, R., and Grosskopf, S. (1997). Productivity and Undesirable Outputs: A Directional Distance Function Approach. *Journal of Environmental Management*, 51(3): 229–240.

Emrouznejad, A., Marra, M., Yang, G. L., and Michali, M. (2023). Eco-efficiency considering NetZero and Data Envelopment Analysis: A critical literature review. *IMA Journal of Management Mathematics*, 34: 599–632.

FAO. (2024). World Food and Agriculture − Statistical Pocketbook 2024. Rome. https://doi.org/10.4060/cd2972en

Fare, R. and Grosskopf, S. (1996). Productivity and intermediate products: A frontier approach. *Economics Letters*, 50(1): 65–70.

Fare, R. and Grosskopf, S. (1997). Intertemporal Production Frontiers: With Dynamic DEA. *Journal of the Operational Research Society*, 48(6): 656.

Fare, R. and Grosskopf, S. (2000). Theory and Application of Directional Distance Functions. *Journal of Productivity Analysis*, 13(2): 93–103.

Fare, R., Grosskopf, S., and Lovell, C. A. K. (1994). *Production Frontiers*. Cambridge University Press.

Fare, R., and Whittaker, G. (1995). An intermediate input model of dairy production using complex survey data. *Journal of Agricultural Economics*, 46(2): 201–213.

Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society*. Series A (General), 120(3): 253.

Fukuyama, H. and Weber, W. L. (2010). A slacks-based inefficiency measure for a two-stage system with bad outputs. *Omega*, 38(5): 398–409.

Fusco, E., Vidoli, F., and Rogge, N. (2020). Spatial directional robust Benefit of the Doubt approach in presence of undesirable output: An application to Italian waste sector. *Omega*, 94: 102053.

Gerami, J., Mozaffari, M. R., Wanke, P. F., and Tan, Y. (2023). Fully fuzzy DEA: A novel additive slacks-based measure model. *Soft Computing*.

Hamidzadeh, F., Mir, S. P., and Naeme, Z. (2024). A novel two-stage network data envelopment analysis model for kidney allocation problem under medical and logistical uncertainty: a real case study. *Health Care Management Science*, 27: 555–579.

Herrera-Restrepo, O., Triantis, K., Trainor, J., Murray-Tuite, P., and Edara, P. (2016). A multi-perspective dynamic network performance efficiency measurement of an evacuation: A dynamic network-DEA approach. *Omega*, 60: 45–59.

Kao, C., and Hwang, S. N. (2021). Measuring the effects of undesirable outputs on the efficiency of production units. *European Journal of Operational Research*, 292(3): 996–1003.

Kao, C., and Hwang, S. N. (2023). Separating the effect of undesirable outputs generation from the inefficiency of desirable outputs production in efficiency measurement. *European Journal of Operational Research*, 311(3): 1097–1102.

Li, F., Wang, Y., Emrouznejad, A., Zhu, Q., and Kou, G. (2022). Allocating a fixed cost across decision-making units with undesirable outputs: A bargaining game approach. *Journal of the Operational Research Society*, 73(10): 2309–2325.

Liang, L., Cook, W. D., and Zhu, J. (2008). DEA models for two-stage processes: Game approach and efficiency decomposition. *Naval Research Logistics*, 55(7): 643–653.

Liu, W., Zhou, Z., Ma, C., Liu, D., and Shen, W. (2015). Two-stage DEA models with undesirable input-intermediate-outputs. *Omega*, 56: 74–87.

Rezaei, A., Dehghan, D. H., Babaei, M. H., and Anvari, A. (2022). Evaluating Environmental Efficiency of European Union Countries with the Approach of Network Data Envelopment Analysis. *Journal of Human, Environment and Health Promotion*, 8(1): 49–61.

Taleb, M., Emrouznejad, A., Charles, V., Khalid, R., and Ramli, R. (2024). An extended-directional mix-efficiency measure: performance evaluation of OECD countries considering NetZero. *Computers and Industrial Engineering*, 189: 109967.

Taleb, N. N., Bar-Yam, Y., and Cirillo, P.(2019). Statistical properties of 401(k) accounts. Physica A: Statistical Mechanics and its Application, 529, 121538.

Tone, K. and Tsutsui, M. (2009). Network DEA: A slacks-based measure approach. *European Journal of Operational Research*, 197(1): 243–252.

Tone, K. and Tsutsui, M. (2010). Dynamic DEA: A slacks-based measure approach. *Omega, 38*(3–4): 145–156.

Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research*, 130(3): 498–509.

Tone, K. (2003). Dealing with undesirable outputs in DEA: A slacks-based measure (SBM) approach. *GRIPS Research Report Series*.

UN Department of Economic and Social Affairs Statistics Division (2024). Statistical Yearbook Sixty-seventh issue. New york: UN Department of Economic and Social Affairs.

.1 appendix A

python program for calculating stage wise and overall efficiency using google collab

```
# Install dependencies
!pip install pulp openpyxl pandas scipy --quiet

import pandas as pd
import numpy as np
from scipy.optimize import linprog
import pulp
from google.colab import files
import io, warnings
warnings.filterwarnings("ignore")

# --------------------
# 1. Upload dataset
# --------------------
print("Upload dataset with columns: DMU, Xc, Xu, Zd, Zud, Zc, Zuc, Yd, Yud")
uploaded = files.upload()
file_name = list(uploaded.keys())[0]

try:
df = pd.read_excel(io.BytesIO(uploaded[file_name]))
except:
df = pd.read_csv(io.BytesIO(uploaded[file_name]))

df.columns = [c.strip().lower() for c in df.columns]
required_cols = ["dmu","xc","xu","zd","zud","zc","zuc","yd","yud"]
if not all(col in df.columns for col in required_cols):
raise ValueError(f"Dataset must have columns: {required_cols}")
```

```
dmus = df["dmu"].astype(str).tolist()
Xc, Xu, Zd, Zud = df["xc"], df["xu"], df["zd"], df["zud"]
Zc, Zuc, Yd, Yud = df["zc"], df["zuc"], df["yd"], df["yud"]

n = len(dmus)

# -------------------
# 2. DDF Solver
# -------------------
def solve_ddf(Xc, Xu, D_out, UD_out):
n = len(Xc)
betas, effs = [], []
for i in range(n):
model = pulp.LpProblem(f"DDF_{i}", pulp.LpMaximize)
beta = pulp.LpVariable("beta", lowBound=0)
lambdas = pulp.LpVariable.dicts("lambda", list(range(n)), lowBound=0)
model += beta
model += pulp.lpSum(lambdas[j]*Xc[j] for j in range(n)) <= Xc[i]*(1-beta)
model += pulp.lpSum(lambdas[j]*Xu[j] for j in range(n)) <= Xu[i]
model += pulp.lpSum(lambdas[j]*D_out[j] for j in range(n)) >= D_out[i]*(1+beta)
model += pulp.lpSum(lambdas[j]*UD_out[j] for j in range(n)) <= UD_out[i]*(1-beta)
model += pulp.lpSum(lambdas[j] for j in range(n)) == 1
model.solve(pulp.PULP_CBC_CMD(msg=0))
b = pulp.value(beta) if pulp.value(beta) else 0
eff = (1-b)/(1+b) if (1+b)!=0 else 0
betas.append(max(min(b,1),0))
effs.append(max(min(eff,1),0))
return betas, effs
```

```
beta1, eff1 = solve_ddf(Xc, Xu, Zd, Zud)
beta2, eff2 = solve_ddf(Zc, Zuc, Yd, Yud)

# -------------------
# 3. SBM Cross-efficiency (multiplier form)
# -------------------
def solve_sbm(Xc, Xu, Y, B):
X_all = np.column_stack([Xc,Xu])
n = len(Xc)
if len(Y.shape)==1: Y=Y[:,None]
if len(B.shape)==1: B=B[:,None]

cross = np.zeros((n,n))
for o in range(n):
c = np.zeros(X_all.shape[1]+B.shape[1]+Y.shape[1])
c[X_all.shape[1]+B.shape[1]:] = -Y[o,:]
A_ub=[]; b_ub=[]
for j in range(n):
row = np.zeros(X_all.shape[1]+B.shape[1]+Y.shape[1])
row[X_all.shape[1]+B.shape[1]:] = Y[j,:]
row[:X_all.shape[1]] = -X_all[j,:]
row[X_all.shape[1]:X_all.shape[1]+B.shape[1]] = -B[j,:]
A_ub.append(row); b_ub.append(0)
A_eq = np.zeros((1,X_all.shape[1]+B.shape[1]+Y.shape[1]))
A_eq[0,:X_all.shape[1]] = X_all[o,:]
A_eq[0,X_all.shape[1]:X_all.shape[1]+B.shape[1]] = B[o,:]
b_eq=[1]
bounds=[(0,None)]*(X_all.shape[1]+B.shape[1]+Y.shape[1])
res=linprog(c,A_ub=np.array(A_ub),b_ub=b_ub,A_eq=A_eq,b_eq=b_eq,bounds=bounds,method
x=res.x if res.success else np.zeros(X_all.shape[1]+B.shape[1]+Y.shape[1])
```

```
v=x[:X_all.shape[1]]; w=x[X_all.shape[1]:X_all.shape[1]+B.shape[1]]; u=x[X_all.shape
for k in range(n):
num=u@Y[k,:]; den=v@X_all[k,:]+w@B[k,:]
cross[k,o]=num/den if den>0 else 0
return np.clip(np.mean(cross,axis=1),1e-6,1)

sbm1 = solve_sbm(Xc.values.reshape(-1,1),Xu.values.reshape(-1,1),Zd.values.reshape(-
sbm2 = solve_sbm(Zc.values.reshape(-1,1),Zuc.values.reshape(-1,1),Yd.values.reshape(

# -------------------
# 4. Ratios & Overall Efficiency
# -------------------
alpha = 0.5  # weight for stage1
ratio1 = np.clip(sbm1/np.array(eff1),1e-6,1)
ratio2 = np.clip(sbm2/np.array(eff2),1e-6,1)
overall = np.clip(alpha*ratio1+(1-alpha)*ratio2,1e-6,1)

# -------------------
# 5. Stage-wise & Overall Rankings
# -------------------
rank_stage1 = ratio1.argsort()[::-1].argsort()+1
rank_stage2 = ratio2.argsort()[::-1].argsort()+1
rank_overall = overall.argsort()[::-1].argsort()+1

results = pd.DataFrame({
 "DMU": dmus,
 "DDF1": eff1, "SBM1": sbm1, "Ratio1": ratio1, "Rank1": rank_stage1,
 "DDF2": eff2, "SBM2": sbm2, "Ratio2": ratio2, "Rank2": rank_stage2,
 "Overall": overall, "Overall_Rank": rank_overall
})
```

```
# -------------------
# 6. Save to Excel
# -------------------
out_file = "Integrated_DDF_SBM_StageOverallRanks.xlsx"
with pd.ExcelWriter(out_file, engine="openpyxl") as writer:
results.to_excel(writer, sheet_name="Results+Ranks", index=False)

print(f"Results + stage & overall ranking saved to {out_file}")
files.download(out_file)
```

.2 appendix B

```
# Install dependencies
!pip install pulp openpyxl pandas scipy matplotlib --quiet

import pandas as pd
import numpy as np
from scipy.optimize import linprog
import pulp
import matplotlib.pyplot as plt
from google.colab import files
import io, warnings
warnings.filterwarnings("ignore")

# -------------------
# 1. Upload Dataset
# -------------------
print("Upload dataset with columns: DMU, Xc, Xu, Zd, Zud, Zc, Zuc, Yd, Yud")
uploaded = files.upload()
```

```
file_name = list(uploaded.keys())[0]

try:
df = pd.read_excel(io.BytesIO(uploaded[file_name]))
except:
df = pd.read_csv(io.BytesIO(uploaded[file_name]))

df.columns = [c.strip().lower() for c in df.columns]
required_cols = ["dmu","xc","xu","zd","zud","zc","zuc","yd","yud"]
if not all(col in df.columns for col in required_cols):
raise ValueError(f"Dataset must have columns: {required_cols}")

dmus = df["dmu"].astype(str).tolist()
Xc, Xu, Zd, Zud = df["xc"], df["xu"], df["zd"], df["zud"]
Zc, Zuc, Yd, Yud = df["zc"], df["zuc"], df["yd"], df["yud"]

n = len(dmus)

# --------------------
# 2. DDF Solver
# --------------------
def solve_ddf(Xc, Xu, D_out, UD_out):
n = len(Xc)
betas, effs = [], []
for i in range(n):
model = pulp.LpProblem(f"DDF_{i}", pulp.LpMaximize)
beta = pulp.LpVariable("beta", lowBound=0)
lambdas = pulp.LpVariable.dicts("lambda", list(range(n)), lowBound=0)
model += beta
model += pulp.lpSum(lambdas[j]*Xc[j] for j in range(n)) <= Xc[i]*(1-beta)
```

```
model += pulp.lpSum(lambdas[j]*Xu[j] for j in range(n)) <= Xu[i]
model += pulp.lpSum(lambdas[j]*D_out[j] for j in range(n)) >= D_out[i]*(1+beta)
model += pulp.lpSum(lambdas[j]*UD_out[j] for j in range(n)) <= UD_out[i]*(1-beta)
model += pulp.lpSum(lambdas[j] for j in range(n)) == 1
model.solve(pulp.PULP_CBC_CMD(msg=0))
b = pulp.value(beta) if pulp.value(beta) else 0
eff = (1-b)/(1+b) if (1+b)!=0 else 0
betas.append(max(min(b,1),0))
effs.append(max(min(eff,1),0))
return betas, effs

beta1, eff1 = solve_ddf(Xc, Xu, Zd, Zud)
beta2, eff2 = solve_ddf(Zc, Zuc, Yd, Yud)

# -------------------
# 3. SBM Cross-efficiency Solver
# -------------------
def solve_sbm(Xc, Xu, Y, B):
X_all = np.column_stack([Xc,Xu])
n = len(Xc)
if len(Y.shape)==1: Y=Y[:,None]
if len(B.shape)==1: B=B[:,None]

cross = np.zeros((n,n))
for o in range(n):
c = np.zeros(X_all.shape[1]+B.shape[1]+Y.shape[1])
c[X_all.shape[1]+B.shape[1]:] = -Y[o,:]
A_ub=[]; b_ub=[]
for j in range(n):
row = np.zeros(X_all.shape[1]+B.shape[1]+Y.shape[1])
```

```
row[X_all.shape[1]+B.shape[1]:] = Y[j,:]
row[:X_all.shape[1]] = -X_all[j,:]
row[X_all.shape[1]:X_all.shape[1]+B.shape[1]] = -B[j,:]
A_ub.append(row); b_ub.append(0)
A_eq = np.zeros((1,X_all.shape[1]+B.shape[1]+Y.shape[1]))
A_eq[0,:X_all.shape[1]] = X_all[o,:]
A_eq[0,X_all.shape[1]:X_all.shape[1]+B.shape[1]] = B[o,:]
b_eq=[1]
bounds=[(0,None)]*(X_all.shape[1]+B.shape[1]+Y.shape[1])
res=linprog(c,A_ub=np.array(A_ub),b_ub=b_ub,A_eq=A_eq,b_eq=b_eq,bounds=bounds,metho
x=res.x if res.success else np.zeros(X_all.shape[1]+B.shape[1]+Y.shape[1])
v=x[:X_all.shape[1]]; w=x[X_all.shape[1]:X_all.shape[1]+B.shape[1]]; u=x[X_all.shap
for k in range(n):
num=u@Y[k,:]; den=v@X_all[k,:]+w@B[k,:]
cross[k,o]=num/den if den>0 else 0
return np.clip(np.mean(cross,axis=1),1e-6,1)

sbm1 = solve_sbm(Xc.values.reshape(-1,1),Xu.values.reshape(-1,1),Zd.values.reshape(
sbm2 = solve_sbm(Zc.values.reshape(-1,1),Zuc.values.reshape(-1,1),Yd.values.reshape

# -------------------
# 4. Ratios & Overall Efficiency
# -------------------
alpha = 0.5
ratio1 = np.clip(sbm1/np.array(eff1),1e-6,1)
ratio2 = np.clip(sbm2/np.array(eff2),1e-6,1)
overall = np.clip(alpha*ratio1+(1-alpha)*ratio2,1e-6,1)

# -------------------
# 5. Rankings
```

```
# -------------------
rank1 = ratio1.argsort()[::-1].argsort()+1
rank2 = ratio2.argsort()[::-1].argsort()+1
rank_overall = overall.argsort()[::-1].argsort()+1

results = pd.DataFrame({
 "DMU": dmus,
 "Stage1_Ratio": ratio1, "Rank1": rank1,
 "Stage2_Ratio": ratio2, "Rank2": rank2,
 "Overall": overall, "Overall_Rank": rank_overall
})

# -------------------
# 6. Save to Excel
# -------------------
out_file = "DDF_SBM_Results_Ranks.xlsx"
with pd.ExcelWriter(out_file, engine="openpyxl") as writer:
results.to_excel(writer, sheet_name="Results+Ranks", index=False)

print(f"Results and rankings saved to {out_file}")
files.download(out_file)

# -------------------
# 7. Bar Plots (Sorted by Overall Rank)
# -------------------
results_sorted = results.sort_values(by="Overall", ascending=False)

dmus_sorted = results_sorted["DMU"]
stage1_eff = results_sorted["Stage1_Ratio"]
stage2_eff = results_sorted["Stage2_Ratio"]
```

```
overall_eff = results_sorted["Overall"]

# ---- Stage 1 & 2 Efficiency Plot ----
x = np.arange(len(dmus_sorted))
width = 0.35
plt.figure(figsize=(15,6))
plt.bar(x - width/2, stage1_eff, width, label="Stage 1")
plt.bar(x + width/2, stage2_eff, width, label="Stage 2")
plt.xticks(x, dmus_sorted, rotation=90)
plt.ylim(0, 1.05)
plt.ylabel("Efficiency")
plt.title("Stage 1 and Stage 2 Efficiencies by DMU (Sorted)")
plt.legend()
plt.tight_layout()
plt.show()

# ---- Overall Efficiency Plot ----
plt.figure(figsize=(15,6))
plt.bar(dmus_sorted, overall_eff, color="green")
plt.xticks(rotation=90)
plt.ylim(0, 1.05)
plt.ylabel("Efficiency")
plt.title("Overall Efficiency by DMU (Sorted)")
plt.tight_layout()
plt.show()
```