MEASURING EFFICIENCY IN EDUCATION USING PISA 2018 DATA

(PISA 2018 VERİLERİNİ KULLANARAK EĞİTİMDE ETKİNLİK ÖLÇÜMÜ)

by

Ece UÇAR, B.S

Thesis

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

in

INDUSTRIAL ENGINEERING

in the

GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

of

GALATASARAY UNIVERSITY

June 2021

This is to certify that the thesis entitled

MEASURING EFFICIENCY IN EDUCATION USING PISA 2018 DATA

prepared by **Ece UÇAR** in partial fulfillment of the requirements for the degree of **Master of Science in Industrial Engineering** at the **Galatasaray University** is approved by the

Examining Committee:

Prof. Dr. E. Ertuğrul KARSAK (Supervisor)
Department of Industrial Engineering
Galatasaray University -------------------------

Doç. Dr. Mehtap DURSUN
Department of Industrial Engineering
Galatasaray University -------------------------

Dr. Öğr. Üyesi Ethem ÇANAKOĞLU
Department of Industrial Engineering
Bahçeşehir University -------------------------

Date: -------------------------

ACKNOWLEDGEMENTS

I gratefully acknowledge the support and guidance of my thesis advisor, Prof. Dr. E. Ertuğrul KARSAK. Without his careful supervision and thoughtful encouragement, this master thesis would never have taken shape. I am so grateful for his contributions to the direction and richness of this research.

My thanks also go out to the faculty members at Galatasaray University, Industrial Engineering Department. It has been a pleasure to be with them since my Bachelor's degree education. I really thank them for shaping my career and giving me valuable professional knowledge.

I extend deepest thanks to my family for their support, love and trust in me. From now on, I will continue to contribute to academic research and create value for public welfare.

Finally, I want to mention that I feel lucky to have completed this master thesis in a healthy way during the difficult pandemic we are going through.

Hope to meet you in healthy and beautiful days.

June 2021
Ece UÇAR

TABLE OF CONTENTS

LIST OF SYMBOLS

PISA : Program for International Student Assessment
OECD : Organization for Economic Co-operation and Development
DEA : Data Envelopment Analysis
OCDE : Organisation de Coopération et de Développement Economique
DMU : Decision-Making Unit
PGB : PISA Governing Board
SDG : Sustainable Development Goal
UNDP : United Nations Development Program
CCR : Charnes, Cooper and Rhodes
BCC : Banker, Charnes and Cooper
VRS : Variable returns-to-scale
TIMSS : Trends in International Mathematics and Science Study
DDF : Directional Distance Function
BoD : Benefit of Doubt
SBM : Slack Base Measure
MCDM : Multi-Criteria Decision-Making
MILP : Mixed Integer Linear Programming
MIP : Mixed Integer Programming

LIST OF TABLES

ABSTRACT

With the need of proposing qualified and impartial education to students from all around the world, the efficient use of educational resources and the measurement of students' success in education have gained increasing importance. In this context, policy makers set significant goals for the sustainable development of education systems and add to their agenda the achievement of quality and equity in education systems. Therefore, proposing reliable indicators that determine how far the goals have been achieved, becomes essential. The PISA (Programme for International Student Assessment) contributes to these efforts by providing valuable information about the success of 15-year-old students in PISA tests and about the factors that influence the education of students. The PISA is a triennial assessment of international education systems and set in cooperation with OECD (Organization for Economic Co-operation and Development) Secretariat. The PISA mainly publishes the test scores of students in reading, mathematics and science and provides some indicators and indexes about the educational resources of students and their social, economic and cultural background.

In this thesis, the PISA 2018 database is used for educational performance assessment of OECD countries. Education process of countries are modelled as systems that produce the success of students in PISA tests by consuming some educational resources. The performance evaluation is conducted by data envelopment analysis (DEA) and common-weight DEA-based models. DEA is a non-parametric technique that measures the relative efficiency of a set of homogenous decision-making units (DMUs) without requiring a priori information about the importance of consumed inputs and produced outputs. The relative efficiency of a DMU is the ratio of the weighted sum of outputs to the weighted sum of inputs. In conventional DEA models, a mathematical program is solved separately for each DMU by maximizing the outputs to inputs ratio. The weights are assigned to outputs and inputs in favour of the evaluated DMU for maximizing its

efficiency score. This situation creates an extreme weight flexibility and the evaluated DMU assigns negligible weights to inputs and/or outputs that represent a poor performance for the related DMU and very high weights to inputs and/or outputs that represent a relatively good performance. Thus, unrealistic weighting schemes are likely to be obtained because of weight flexibility and result in impractical evaluations.

The DEA model dichotomizes the DMUs as efficient and inefficient by assigning an efficiency score of one to efficient units and an efficiency score less than one to inefficient units. However, a further discrimination among the efficient units is a need for full ranking of DMUs. As mentioned above, conventional DEA models have some limitations such as weight flexibility creating unrealistic weighting schemes and insufficient discriminatory characteristics. In order to overcome these shortcomings, common-weight DEA-based models can be considered. These models measure the efficiency scores of DMUs by using a common set of weights, provide computational savings and enhance the discriminating power of the analysis. Despite these advantages, to the best of our knowledge, the common-weight DEA-based models have not been used yet in evaluating educational efficiency using PISA data. Therefore, this thesis aims to provide a comparative analysis using a number of selected common-weight DEA-based approaches for educational performance assessment utilizing the most up-to-date edition of the PISA database. The outputs of the analysis are reading, mathematics and science scores of students in PISA tests whereas the inputs are the number of teachers per number of students, total learning time in minutes per week and average number of computers available in schools for education purposes. The efficiency assessment is conducted for 29 OECD countries.

The results obtained by different models are compared in terms of efficiency scores, rankings and weight dispersions. In majority of the models, Estonia and Finland are determined as the most efficient countries. The similarities between the sets of rankings obtained by different models are tested with Spearman rank correlation. While the conventional DEA model identifies ten countries as efficient, the majority of applied common-weight DEA-based models achieve full discrimination among countries.

This assessment can be a good reference for policy-makers to observe the current situation of their countries in terms of education efficiency and to derive political implications for the improvement of their education systems. Furthermore, as this efficiency analysis provides an international comparison of OECD countries' education systems, it provides a benchmarking opportunity among these countries. Moreover, as the PISA 2018 is the latest edition of the PISA cycle, the efficiency assessment provides a more up-to-date analysis for OECD countries.

RÉSUMÉ

Avec la nécessité de proposer une éducation qualifiée et impartiale aux étudiants du monde entier, l'utilisation efficace des ressources éducatives et la réussite des étudiants dans l'éducation ont pris une importance croissante. Dans ce contexte, les décideurs politiques fixent des objectifs importants pour le développement durable des systèmes éducatifs et ajoutent à leur programme la réalisation de la qualité et de l'équité dans les systèmes éducatifs. Par conséquent, il devient essentiel de proposer des indicateurs fiables qui déterminent dans quelle mesure ces objectifs ont été atteints. Le PISA (Programme pour Evaluation Internationale des Etudiants) contribue à ces efforts en fournissant des informations précieuses sur la réussite des étudiants aux tests PISA et sur les facteurs qui influencent l'éducation des étudiants. Le PISA est une évaluation triennale des systèmes éducatifs internationaux et établie en coopération avec le Secrétariat de l'OCDE (Organisation de Coopération et de Développement Economique). Le PISA publie principalement les résultats de tests des élèves en lecture, en mathématiques et en sciences et fournit des indicateurs sur les ressources éducatives des élèves et leur origine sociale, économique et culturelle.

Dans cette thèse, la base de données PISA 2018 est utilisée pour l'évaluation des performances éducatives des pays de l'OCDE. Les processus éducatifs des pays sont modélisés comme des systèmes qui assurent la réussite des élèves aux tests PISA en consommant certaines ressources pédagogiques. L'évaluation des performances appliquée dans cette thèse est réalisée par des modèles d'analyse d'enveloppement de données (DEA). La DEA est une technique non paramétrique qui mesure l'efficience relative d'un ensemble d'unités de décision homogènes (DMU) sans nécessiter d'informations a priori sur l'importance des intrants consommés et des extrants produits. L'efficience relative d'une DMU est définie comme le rapport entre la somme pondérée des extrants et la somme pondérée des intrants. Dans les modèles DEA conventionnels,

un programme mathématique est résolu séparément pour chaque DMU en maximisant le rapport extrants pondérés / intrants pondérés. Les poids sont attribués aux extrants et aux intrants en faveur de la DMU évaluée pour maximiser son score d'efficience. Cette situation crée une flexibilité de poids extrême et la DMU évaluée attribue des pondérations négligeables aux indicateurs qui montrent une mauvaise performance pour la DMU associée et des pondérations très élevées aux indicateurs qui montrent une performance relativement bonne. Ainsi, des schémas de pondération irréalistes peuvent être obtenus en raison de la flexibilité du poids et aboutir à des évaluations peu pratiques.

Le modèle DEA dichotomise les DMUs comme étant efficients et inefficients en attribuant un score d'efficience d'un aux unités efficients et un score d'efficience inférieur à un aux unités inefficients. Cependant, une discrimination entre les unités efficientes est la nécessité d'un classement complet des DMUs. Comme mentionné ci-dessus, les modèles DEA conventionnels ont certaines limitations telles que la flexibilité de poids créant des schémas de pondération irréalistes et des caractéristiques discriminatoires insuffisantes. Afin d'éviter ces problèmes, des modèles basés sur la DEA à pondération commune peuvent être envisagés. Ces modèles mesurent les scores d'efficience des DMUs en utilisant un ensemble commun de poids, fournissent des économies de calcul et améliorent le pouvoir discriminant de l'analyse. Malgré ces avantages, à notre connaissance, les modèles basés sur la DEA à pondération commune n'ont pas été encore utilisés pour évaluer l'efficience de l'éducation à l'aide des données PISA. Par conséquent, cette thèse vise à fournir une analyse comparative en utilisant un certain nombre d'approches basées sur la DEA à pondération commune pour l'évaluation des performances éducatives en utilisant l'édition la plus à jour de la base de données PISA. Les données de sortie des modèles DEA utilisés sont les scores en lecture, en mathématiques et en sciences des élèves aux tests PISA et les données d'entrée sont le nombre d'enseignants par nombre d'élèves, le temps d'apprentissage total en minutes par semaine et le nombre moyen d'ordinateurs disponibles dans les écoles à des objectives éducatives. L'évaluation de l'efficience est réalisée pour 29 pays de l'OCDE.

Les résultats obtenus par différents modèles sont comparés en termes de scores d'efficience, de classements et de dispersions de poids des intrants et extrants. Dans la majorité des modèles, l'Estonie et la Finlande sont sélectionnées comme les pays meilleurs en terme d'éducation. Les similitudes entre les ensembles de classements obtenus par les différents modèles sont testées avec le coefficient de corrélation de rang de Spearman. Alors que les modèles DEA conventionnels identifient dix pays comme étant efficients, la majorité des modèles basés sur la DEA à pondération commune réalisent une discrimination totale entre les pays.

Cette évaluation peut être une bonne référence pour les décideurs politiques afin de voir la situation actuelle de leurs pays en termes d'efficience de l'éducation et pour en tirer des implications politiques pour l'amélioration de leurs systèmes éducatifs. Grâce à cette analyse, un classement international des systèmes éducatifs en considérant leur efficience est obtenu. De plus, comme le PISA 2018 est la dernière édition du cycle PISA, l'évaluation de l'efficience fournit une analyse très récente pour les pays de l'OCDE.

ÖZET

Dünya'nın dört bir yanındaki öğrencilere nitelikli ve tarafsız eğitim sunmanın gerekliliğinden dolayı, eğitim kaynaklarının etkin kullanımı ve öğrencilerin eğitimdeki başarısının ölçülmesi giderek artan bir önem kazanmıştır. Bu bağlamda eğitimden sorumlu yöneticiler, eğitim sistemlerinin sürdürülebilir gelişimi için önemli hedefler belirlemekte ve eğitim sistemlerinde kalite ve eşitliğin sağlanmasını gündemlerine eklemektedir. Bu nedenle, hedeflere ne kadar ulaşıldığını belirleyen güvenilir göstergelerin üretilmesi zorunlu hale gelmiştir. PISA (Uluslararası Öğrenci Değerlendirme Programı), 15 yaşındaki öğrencilerin PISA testlerindeki başarıları ve öğrencilerin eğitimini etkileyen faktörler hakkında değerli bilgiler sağlayarak bu çabalara katkıda bulunmaktadır. PISA veri tabanı, uluslararası eğitim sistemlerinin üç yıllık bir değerlendirmesidir ve OECD (Ekonomik İşbirliği ve Kalkınma Teşkilatı) Sekreterliği ile işbirliği içinde oluşturulmaktadır. PISA genel olarak öğrencilerin okuma, matematik ve fen bilimlerindeki test puanlarını yayınlamaktadır; bununla beraber öğrencilerin erişebildiği eğitim kaynakları, sosyal, ekonomik ve kültürel geçmişleri hakkında bazı göstergeler ve indeksler üretmektedir.

Bu yüksek lisans tezinde, OECD ülkelerinin eğitim performanslarının değerlendirmesi için PISA 2018 veri tabanı kullanılmıştır. Ülkelerin eğitim süreci, bazı eğitim kaynaklarını tüketerek PISA testlerinde öğrencilerin başarısını üreten sistemler olarak modellenmiştir. Performans değerlendirmesi, veri zarflama analizi (VZA) modelleri ile yapılmıştır. VZA, kullanılan girdilerin ve üretilen çıktıların önem derecesi hakkında ön bilgi gerektirmeden bir dizi homojen karar verme biriminin (KVB'ler) göreceli etkinliğini ölçen parametrik olmayan bir tekniktir. Bir KVB'nin göreceli etkinlik derecesi, çıktıların ağırlıklı toplamının girdilerin ağırlıklı toplamına oranı olarak tanımlanmaktadır. Geleneksel VZA modellerinde, ağırlıklı çıktıların ağırlıklı girdilere oranını maksimize ederek her KVB için ayrı bir matematiksel programlama modeli çözülmektedir.

Girdi ve çıktıların ağırlıkları, etkinlik puanını en büyüklemek için değerlendirilen KVB lehine belirlenmektedir. Bu durum, ağırlıkları belirlerken büyük bir esneklik oluşturmaktadır ve değerlendirilen KVB, kötü performans gösterdiği girdi ve/veya çıktılara önemsiz ağırlıklar ve nispeten iyi performans gösterdiği girdi ve/veya çıktılara çok yüksek ağırlıklar atamaktadır. Bu nedenle, ağırlık esnekliği sonucunda gerçekçi olmayan ağırlıklandırma söz konusu olabilmekte ve pratik olmayan, gerçeği yansıtmayan değerlendirmeler ortaya konmaktadır. İdeal bir etkinlik değerlendirmesinde, ilgili KVB'nin performansını etkileyen girdi ve çıktıların tamamının hesaplamaya katılması daha gerçekçi sonuçlar doğurmaktadır.

Klasik VZA modeli, etkin birimlere bire eşit olan etkinlik puanını, etkin olmayan karar verme birimlerine ise birden düşük etkinlik puanlarını atayarak KVB'leri etkin ve etkin olmayan olarak ikiye ayırmaktadır. Çoğu analizde, etkinlik karşılaştırmasında KVB'lerin tam sıralamasına ihtiyaç duyulmaktadır ve dolayısıyla klasik VZA modelleri yetersiz kalmaktadır. Yukarıda bahsedildiği gibi, geleneksel VZA modelleri, gerçekçi olmayan sonuçlar yaratan ağırlık esnekliği ve yetersiz ayrımcı özellikler gibi bazı sınırlamalara sahiptir. Bu eksikliklerden kaçınmak için, ortak ağırlıklı VZA tabanlı modellere yönelmek uygundur. Bu modeller, ortak bir ağırlık seti kullanarak KVB'lerin etkinlik puanlarını ölçmektedir. Genellikle bir ya da iki matematiksel programlama modeliyle tüm ağırlıklar ve etkinlik puanları hesaplanabilmekte ve daha da önemlisi VZA'ya göre etkin birimler arasında ayırt etme olanağı sağlanabilmektedir. Bu avantajlara rağmen, bildiğimiz kadarıyla, ortak ağırlıklı VZA tabanlı modeller, PISA verileri kullanılarak eğitim performansını değerlendirmede henüz kullanılmamıştır. Bu nedenle, bu tez çalışması, PISA veri tabanının en güncel versiyonunu kullanarak ülkelerin eğitim performansının değerlendirmesi için literatürde yer alan ortak ağırlıklı VZA temelli yaklaşımlarını kullanarak karşılaştırmalı bir analiz yapmayı amaçlamaktadır. Tez çalışmasında kullanılan VZA modellerinin çıktıları, öğrencilerin PISA sınavlarındaki okuma, matematik ve fen bilimleri puanlarıdır; diğer yandan girdiler ise öğrenci sayısı başına düşen öğretmen sayısı, haftalık dakika cinsinden toplam öğrenme süresi ve eğitim amaçlı okullarda bulunan ortalama bilgisayar sayısıdır. Etkinlik değerlendirmesi 29 OECD ülkesi için yapılmıştır.

Farklı modellerle elde edilen sonuçlar, etkinlik puanları, sıralamaları ile girdi ve çıktıların ağırlık dağılımları açısından karşılaştırılmıştır. Modellerin çoğunda, Estonya ve Finlandiya eğitim açısından en etkin ülkeler olarak belirmektedir. Farklı modellerle elde edilen sıralama setleri arasındaki benzerlikler, Spearman sıra korelasyonu ile test edilmiştir. Geleneksel VZA modelleri on ülkeyi etkin olarak tanımlarken, uygulanan ortak ağırlıklı VZA tabanlı modellerin çoğu, ülkeler arasında tam bir sıralama sağlamaktadır.

Bu tez çalışmasıyla elde edilen eğitim etkinliği değerlendirmesi, eğitimden sorumlu yöneticilerin ve devlet yetkililerinin ülkelerinin mevcut durumunu algılamalarına ve eğitim sistemlerinin iyileştirilmesi için çıkarımlar elde etmelerine olanak sağlamaktadır. Üstelik bu çalışma sayesinde 29 farklı OECD ülkesinin, eğitim etkinliği açısından uluslararası ölçekte bulunduğu konuma ulaşılmaktadır. Ayrıca, PISA 2018, PISA döngüsünün en son sürümü olduğundan, yapılan eğitim sistemi etkinliği değerlendirmesi OECD ülkeleri için en güncel analiz olarak belirmektedir.

1. INTRODUCTION

Governments are looking for international comparisons of education opportunities and systems in order to develop policies that enhance efficiency in students' social and economic prospects, schooling, and the use of educational resources to meet increasing demands. The PISA (Program for International Student Assessment) contributes to these efforts by measuring students' ability to use their reading, mathematics, and science knowledge and skills to meet real-life challenges. Together with OECD (Organization for Economic Co-operation and Development) country policy reviews, PISA results can be used to assist governments in building more efficient and equitable education systems. In this way, measuring efficiency in education is a need for fixing the inefficient components of the education systems and moving toward a more efficient and sustainable future for students' education.

In this context, this study aims to evaluate the educational efficiency of OECD countries using PISA 2018 database to compare the current situation of the education systems and to derive political implications for the efficient use of the educational resources. For this purpose, performance evaluation will be conducted by models based on Data Envelopment Analysis (DEA). Since the 1990s, efficiency evaluation in education is treated as an attractive area of research. Authors assessed efficiencies with different levels of analysis (course-level, school-level, and country-level) and with different variables related to students, families and education institutions. According to a recent literature survey regarding efficiency in education (Witte & López-torres, 2017), DEA is the most widely used methodology for educational performance assessment.

DEA is a decision making tool based on mathematical programming, which is used for assessing relative efficiency of homogenous decision making units (DMUs) that consume multiple inputs to produce multiple outputs. Traditional DEA models identify efficient and inefficient units without requiring a priori information about the importance of inputs and outputs (Karsak & Ahiska, 2005). Despite their widespread use as a decision tool, these models also have a number of limitations. First, conventional DEA models are solved *n* times to measure the efficiency scores of all DMUs, where *n* is the number of homogenous DMUs to be evaluated. Second, the DMUs are not evaluated on a common basis for input and output weights because each DMU is allowed to choose its own weights for maximization of its relative efficiency score. The weight flexibility in DEA may result in a DMU to appear efficient by weighting an input and/or output impractically high while assigning negligible weights to other inputs and/or outputs (Karsak & Ahiska, 2007). This situation leads to unrealistic weighting schemes and impractical evaluations. Third, while these models dichotomize the DMUs as “efficient” (with an efficiency score of 1) and “inefficient” (with an efficiency score less than 1), they cannot provide a further ranking among the efficient DMUs. Therefore, the discriminating power of these models is not sufficient either for full ranking of DMUs or for identifying the best DMU among the DMUs whose efficiency scores are equal to unity. To overcome this issue, the super efficiency measure was proposed (Andersen and Petersen, 1993). This measure enables to define super-efficient units with efficiency scores greater than one by excluding the constraint that limits the efficiency score of the evaluated DMU to unity. The super efficiency model improves the discriminating power but it has some limitations such as weight flexibility, obtaining very high efficiency scores and infeasibility of solutions. Another measure that is proposed to control weight flexibility is weight restrictions that are added as constraints to the DEA model in order to avoid assigning zero-weights to inputs and outputs. Weight restrictions are also used to incorporate a priori information on the importance of inputs and outputs (Allen *et al.*, 1997). However, this approach contradicts with the objective-weighting scheme by manipulating input & output weights. Another measure that is proposed to improve discriminating power of the DEA model is the cross-efficiency analysis (Doyle & Green, 1994). For this analysis, a peer appraisal matrix is constructed where a DMU is evaluated according to weights chosen in favor of other DMUs. Despite improving the discriminatory characteristics, cross-efficiency measure requires solving considerably greater number of formulations and increases computational burden.

In order to avoid the shortcomings of conventional DEA models, common-weight DEA-based models can be considered. These models measure the efficiency scores of DMUs by using a common set of weights, provide computational savings and enhance the discriminating power of the analysis. Despite these advantages, to the best of our knowledge, the common-weight DEA-based models have not been used in evaluating educational efficiency using PISA data. Therefore, this thesis aims to provide a comparative analysis using a number of selected common-weight DEA-based approaches for educational performance assessment utilizing the most up-to-date edition of the PISA database.

The rest of the thesis is organized as follows. Section 2 presents the basics of PISA 2018 database, while section 3 provides a review of the literature about DEA models applied to PISA data and recently proposed common-weight DEA-based models. Section 4 presents the methodological approach. In Section 5, comparative evaluation of OECD countries using PISA 2018 data is illustrated via common-weight DEA-based models and a novel approach is proposed for comparing results. Discussion and concluding remarks are outlined in the last section.

2. PISA DATA

The Program for International Student Assessment (PISA) conducted by OECD, examines the knowledge and skills acquired by 15-year-old students near the end of their compulsory education. The triennial assessment, launched in 1997, not only assesses the students' level of knowledge but it also examines how well students can integrate the acquired knowledge and skills in daily life routines and can apply what they have learned in unfamiliar situations, both in and outside of school (OECD, 2017). The assessment has three core subjects such as reading, mathematics and science literacy. In addition, it includes optional assessments such as financial literacy and global competence according to the preference of participating countries (OECD, 2019c).

In each PISA cycle, one of the core subjects is chosen as the major domain and tested in details. For PISA 2018 cycle, the major domain was reading, as it was in 2000 and 2009. Beside the assessment of core domains, additional information about students' backgrounds, schools, teachers and parents is collected via questionnaires. The program combines information collected by questionnaires and information about proficiency level of students on core subjects and then provides three main types of outcomes. First, it makes provision about the basic indicators of knowledge and skills of students. Second, it supplies information about the background indicators and enables to relate the students' knowledge and skills with numerous demographic, social, economic and educational variables. Third, as it is a triennial assessment, it provides insights on trends and changes in PISA test scores and evaluates the links of these changes with student-level, school-level and system-level background indicators and outcomes (OECD, 2019b).

Policy-makers around the world understand the relative strengths and weaknesses of the education systems of their countries through PISA findings. Additionally, they find the opportunity to compare the PISA results with other countries and form benchmarks for improvements in the educational outcomes of their countries. For this comparison, PISA database is valuable and very rich as the PISA 2018 results represent 31 million 15-year-old students from the participation of 710 000 students in the schools of the 79 participating countries and economies including 37 OECD countries (OECD, 2019b). PISA results are interpreted by using a scale. On this scale, the average PISA score among OECD countries is 500 with a standard deviation of 100, which means students scored between 400 and 600 for core subjects such as mathematics, reading, and science (OECD, 2017).

The assessment process for reading, mathematics and science knowledge and skills start with computer-based tests during two hours. Participating countries that declare the lack of technical equipment can also take the test in a paper-based format. Test items are a mixture of multiple-choice and open-ended questions and different item combinations are made for each student (OECD, 2019b). In addition to computer-based tests, student questionnaires and school questionnaires are also conducted in order to collect information about students themselves, their homes, socio-economic backgrounds, educational experiences, school systems and educational resources. As an option, some countries also conduct questionnaires for parents and teachers about their background. Countries/economies can also use questionnaires for collecting information about students' habit of information technology use, their future career, well-being and life satisfaction (OECD, 2017).

Some important features make PISA a unique assessment mechanism. It is seen as the most comprehensive and accurate international education assessment program by OECD policy-makers (OECD, 2019b). Because of the quality-assurance mechanisms applied in translation, sampling and data collection phases, PISA results have a high degree of validity and reliability. PISA's unique features are listed as policy orientation, innovative concept of literacy, relevance to life-long learning, regularity and breadth of coverage (OECD, 2019b). Policy orientation refers to the relationships established between PISA results and background variables. Insights can be derived from these relationships for educational policy improvements. Innovative concept of literacy points out the understanding of literacy by PISA experts. For PISA, literacy means the student's ability of applying knowledge and skills in real-life challenges. For example, reading literacy defines a student's capacity to understand and reflect on texts for achieving a goal and communicating in the society. Mathematical literacy defines a student's ability to use mathematical reasoning for explaining some phenomena. Finally, scientific literacy defines a student's proficiency in reasoning scientifically by utilizing data and technology. Moreover, relevance to life-long learning is tested with the PISA questions that are asked to ascertain student's motivation, strategies and learning objectives. Finally, regularity and breadth of coverage are also among the key features of the PISA as it is a cycle applied every 3-years and covers all 37 OECD countries and 42 partner countries (OECD, 2019c).

It is important to notice that PISA is the result of a collaborative project between OECD and partner governments. The evaluations are conducted co-operatively, approved by participating countries and implemented by national organizations. The members of this co-operation are mainly the PISA Governing Board (PGB) and National Project Managers of PISA. The PGB members are the senior level representatives from all participating countries and are responsible for the policy orientation of PISA. National Project Managers ensure the accurate implementation of PISA at national-level in each participating country. Finally, the OECD secretariat has the responsibility for the management of the program. The secretariat communicates with PGB members and National Project Managers and monitors the implementation of the program on a daily basis. Furthermore, OECD is also responsible for official report publications and managerial implications (OECD, 2019b).

Nowadays, governments are increasingly looking to develop education policies by making international comparisons of education systems. Their objective is to enhance the establishment of more effective and equitable education systems in the context of Sustainable Development Goal 4 (SDG 4). The SDG 4 set by the UNDP (United Nations Development Program) in September 2015, aims to ensure "inclusive and equitable quality education and promote lifelong learning opportunities for all" (OECD, 2019). The goal is detailed in terms of targets, and global and thematic indicators designed to measure the achievement of the goal. OECD is responsible for developing data about these indicators that are representative for the SDG 4. In this context, the PISA is one of the important tools that contributes to these efforts by promoting indicators and thus measuring the progress towards the achievement of SDG 4. For example, the proficiency levels in reading scores of PISA are representatives of quality in education, which are among important points of SDG 4 (OECD, 2019b).

Finally yet importantly, PISA database is a very important source for researchers considering the indicators, test scores and indexes that it provides. PISA data are published in details and open for further analysis of researchers. Technical reports and the online database published by PISA contain all the information about students' results, answers and comparisons of participating countries in different categories of reading, mathematics and science. Thus, this thesis uses PISA data for efficiency measurement and performance comparison on a country-level. PISA data are considerable representatives of current situation of education systems of OECD countries as all world leaders and OECD members accept it as a reliable source.

3. LITERATURE REVIEW

3.1 DEA Applications in Education

Since 1990s, there is an effort to measure the efficiency in education at different levels such as class, course, school or country. According to Witte & López-Torres (2017), data envelopment analysis (DEA) is the most frequently used methodology throughout these studies and the PISA dataset is a valuable source for input and output data.

A well-known conventional DEA model is the CCR model proposed by Charnes, Cooper and Rhodes (1978). Although the CCR model identifies efficient DMUs, it cannot rank the efficient units as all the efficient units possess an efficiency score of 1. Moreover, the model is solved separately for each DMU and allows the evaluated DMU to select the weights for maximizing its efficiency. Therefore, it cannot provide a common weight basis for an overall assessment and, in general, it cannot identify the best DMU.

The CCR model is used in the context of education efficiency and PISA data with the output-orientation (Yalçin & Tavşancil, 2014; Aparicio *et al.*, 2020) or with the input-orientation. Marti Selva & Medina (2018) applied the CCR model for education efficiency comparison of European and Asian countries. The data source of inputs and outputs was the results for TIMSS 2015 (Trends in International Mathematics and Science Study). The applied methodology separated the efficient countries from the non-efficient ones, however further discrimination among the efficient units could not be achieved. Margaritis *et al.* (2021) investigated relative efficiency of upper secondary schools in the region of Central Greece by using the input-oriented CCR model and BCC model. Malmquist Indexes for years 2015-2018 were also calculated. Data source for the input and output indicators was the records of upper secondary schools and regional directorate of Central Greece.

When DEA applications in education where the PISA dataset is used as the source of inputs and outputs are analyzed, the most frequently used DEA model is the BCC model (Banker, Charnes and Cooper, 1984). This model is one of the conventional DEA models and differs from the CCR model (Charnes, Cooper and Rhodes, 1978) with its assumption of variable-returns to scale (VRS). The BCC model is solved *n* times, where *n* is the number of DMUs, and each time it generates weights for the inputs and outputs in favor of the evaluated DMU. This situation leads to computational disadvantages and unrealistic weighting schemes. Moreover, the BCC model identifies the DMUs with an efficiency score of one as BCC-efficient but cannot allow further discrimination among these efficient units.

The PISA studies using the BCC model can be classified in three groups. The first group used the classical BCC model (Santín & Sicilia, 2015; Segovia-Gonzalez *et al.*, 2020). The second group applied the BCC model with an output-orientation (Afonso & St. Aubyn, 2006; Mancebón *et al.*, 2012; Coco & Lagravinese, 2014; Agasisti, 2014; Agasisti & Zoido, 2019; Deutsch *et al.* 2019) . The last group applied the BCC model with an input-orientation (Bogetoft *et al.*, 2015; Mou *et al.*, 2019). Although these models suffer from the shortcomings of the BCC model such as insufficient discriminating power, computational disadvantage and weight flexibility, the use of output-oriented or input-oriented DEA models have several advantages. The output-oriented BCC model can answer the question of how much the outputs can be improved for better performance and the input-oriented DEA model answers the question of how much the inputs can be reduced to have enhanced performance.

Among the studies using output-oriented BCC models, one group is distinguished based on their two-stage approach (Afonso & St. Aubyn, 2006; Coco & Lagravinese, 2014; Agasisti, 2014; Agasisti & Zoido, 2019). In this group of studies, they firstly applied the DEA model to measure the efficiency scores, secondly they applied a regression model to explain the efficiency scores, in other words to identify and represent variables that are associated with PISA test performances. To illustrate, Coco & Lagravinese (2014) used the output-oriented DEA model under VRS to evaluate efficiency scores of OECD countries for the first stage of the study. Cumulative spending in education (in USD) was used as the input; mathematics, reading and science scores were used as the outputs from the PISA 2009 database. Then, the efficiency scores were corrected with a bootstrap procedure introduced by Simar & Wilson (2007) and explained in a truncated regression with independent variables such as parents' educational attainment, immigrant status, labor-market indicators, cronyism, the number of students per class, teaching time, examination and inspections.

Among the studies using PISA data, there are many other DEA approaches applied for different purposes. For example, a more recent study proposed a new approach based on the weighted additive DEA model with the assumption of variable-returns to scale for the efficiency analysis of Spanish schools (Aparicio, Cordero and Pastor, 2017). The weighted additive model evaluates input-saving and output-expansion at the same time, but it needs to be solved *n* times and generates a different weighting scheme for each DMU. The problem required with the resolution of 902 optimization problems in 3 hours. The results of the study showed that the schools could increase their output levels by 12% for better performance. However, the weight dispersion for inputs and outputs as well as discriminatory characteristics of the model were not elaborated in this study.

Another study analyzed the PISA 2012 database to compare the education performance of 34 countries with the non-radial DEA model and the Tobit regression analysis to find different levels of inefficiency in the outputs (Aparicio *et al.*, 2018). The non-radial approach allows evaluating the schools whose concern is to improve one dimension of the educational output greater than the other dimension. The results provided a full ranking between the countries but the model did not use a common-weight basis. Thus, an objective evaluation could not be achieved because each DMU prioritized the weights of inputs/outputs that showed a superior performance for the evaluated DMU.

Shamohammadi and Oh (2019) examined the efficiency scores of private universities in Korea by applying a two-stage network DEA model. They modelled the efficiency of universities with two stages as teaching and research. The efficiencies of universities were measured with the feedback model implemented between teaching and research stages by using network DEA. The results showed improved discrimination characteristics in comparison to the traditional DEA model. However, the weight dispersion of the inputs and outputs were not interpreted in the study.

All of the studies mentioned above used precise inputs and outputs. However, a fuzzy DEA approach was also applied in the context of PISA 2015 database in order to examine relative performance of schools from the United States (Aparicio *et al.,* 2019). The authors evaluated the socio-economic status of students as the fuzzy input, and mathematics and reading test scores as the fuzzy outputs. The study enabled to evaluate the impact of imprecise data on educational efficiency scores. However, the study did not provide a common assessment among the DMUs since the model was solved separately for each DMU. The results of fuzzy DEA model showed better discrimination than the traditional DEA model by ranking 94% of BCC-efficient units.

Foladi *et al.* (2020) applied dynamic DEA and inverse DEA for the performance evaluation of university faculties. Dynamic DEA was used to add the time component to the model and inverse DEA was used to interpret changes on input and output levels. According to the results, dynamic DEA provided more discrimination than static (conventional) DEA. As the focus of the study was on the change of efficiency during time, weight dispersion for inputs and outputs was not considered.

The Directional Distance Function (DDF) approach was also applied to PISA data with DEA models. Ben Yahia *et al.* (2018) examined Tunisian education efficiency based on a DDF and DEA methodology. Rebai *et al.* (2020) estimated the DDF with DEA and applied regression as a machine learning approach for the efficiency assessment of Tunisian schools. The learning times in science, language, and mathematics were the inputs from PISA 2012 database and the drop-rate was the undesirable output. The DDF approach was used to incorporate the undesirable output to the efficiency analysis. As the model was solved separately for each DMU, it lacks simultaneous assessment with common weights. The results of the efficiency analysis showed that 22% of 105 Tunisian schools were efficient, but the best performing school could not be identified with the DDF approach.

More recently, Coco *et al*. (2020) computed the relative education efficiency of OECD countries with the Benefit of Doubt (BoD) approach known as the DEA without input and compared the performances between PISA 2006, 2009, 2012 and 2015 cycles. The outputs included mathematics, reading and science scores and the model included a dummy input. A full discrimination of efficiency scores could not be achieved with the BoD approach. For the BoD approach, a composite index of efficiency was estimated based on the combination of weights, which was the more convenient for the evaluated school. In consequence, the BoD approach shows a subjective evaluation of weights for each DMU. Lagravinese *et al.* (2020) applied DEA in Slack Base Measure (SBM) version in order to compare the performances of 60 different countries utilizing PISA 2009 and 2012. The outputs included the mathematics, language and science scores and the learning time (per week in minutes) was the only input. The SBM model evaluated the effect of socio-economic and cultural status of students on outputs and was solved separately for each DMU, thus could not provide a common assessment of weights. Moreover, the best performing DMU could not be identified.

In order to examine the efficiency of higher education in Mexico, a dynamic-network DEA model was applied for the period between 2008 and 2016 (Delfin-Ortega & Navarro-Chávez, 2020). The model allowed assessing the time component of efficiency with carryover factors between different periods but the higher education institutions were classified as efficient with an efficiency score of 1 and inefficient with an efficiency score less than 1. Thus, the model could not help to allow a further discrimination among the efficient higher education units.

A non-radial super efficiency DEA framework was proposed to measure the research efficiency of different disciplines at Chinese Universities (Su *et al*., 2020). They demonstrated that the model has better discriminatory characteristics than traditional DEA methods, solved the infeasibility problem of super-efficiency model. Furthermore, they calculated common-weights for the inputs and outputs by using a DEA-MCDM approach (H. Wang, 2015) and obtained objective set of weights, thus avoided subjective preference of DMUs. Their approach has several advantages such as improved discriminatory characteristics and common-weight basis but as they divided their approach into different steps (first, finding weights and then calculating efficiency scores with another model); it can be complex in terms of implementation and infeasibility.

3.2 Common-weight DEA-based Models

Recently, several researchers have worked on the literature on common weight DEA-based approaches. Some of them proposed models on finding the best decision-making unit (DMU) while others have focused on providing a full ranking between the DMUs.

Karsak and Ahiska (2007) developed a minimax efficiency model for determining the best DMU taking into consideration multiple inputs and outputs. The discriminating power of the approach was proved with examples mentioned in earlier researches and the model achieved a better weight dispersion for inputs and outputs.

Foroughi (2011) proposed a mixed integer linear programming (MILP) model for finding the most efficient unit and extended the model for ranking all extreme efficient units. The model was tested with real-life examples such as facility layout design selections.

Wang and Jiang (2012) proposed three alternative MILP models for identifying the most efficient DMU under different returns to scale. The proposed MILP models were illustrated with four numerical examples.

Sun et al. (2013) developed two different common-weight DEA models considering ideal and anti-ideal DMU for performance evaluation and full ranking. The models were applied for the performance assessment of Asian lead frame firms and flexible manufacturing systems.

Lam (2015) introduced a new MILP model whose objective was similar to the super-efficiency model. The proposed model was compared with two recent models with two illustrative examples about facility layout design and hospital management.

Toloo (2015) developed a new minimax MILP model for finding the most efficient DMU. The approach illustrated with three different case studies from the literature and the results were compared with other methods.

Carrillo and Jorge (2016) constructed a multi-objective DEA approach for ranking alternatives. They used the concept of distance to an ideal. Therefore, the model chose the common set of weights by minimizing the total Tchebychev distance or Manhattan distance of DMUs to an ideal point.

Toloo and Salahi (2018) presented a new nonlinear mixed integer-programming (MIP) model whose discriminating power was high enough for fully ranking all DMUs. The model was linearized and two examples were used from the literature for the comparison with other models.

Karsak and Goker (2020) proposed a novel common weight multiple criteria decision making (MCDM) approach based on the model proposed by Karsak and Ahiska (2007) for the selection of best performing DMU. The approach was compared with other methods and applied to two case studies. The results showed that the model provided an enhanced discriminating power for the rank-order and improved weight dispersion for inputs and outputs.

Lately, Özsoy et al. (2021) proposed a new MIP model without epsilon based on the approach proposed by Toloo and Salahi (2018). Discriminating power of the model was illustrated with a simulation study where it was shown that the model was able to provide full ranking.

For more details, Table 3.2.1 provides advantages and limitations of the developed methodologies in a tabular format.

Table 3.2.1: Advantages and Limitations of Common-weight DEA-based Approaches

Author(s)	Year	Advantages	Limitations
Karsak and Ahiska	2007	–Extends the model proposed by Karsak and Ahiska (2005) by considering multiple inputs in efficiency analysis – Enhanced discriminatory characteristics compared with conventional DEA	– Requires a subjective evaluation for determining the value of *k* – May find inconsistent results with the first stage of the model
Foroughi	2011	-Selects the best DMU in one step	-Needs to specify an assurance region for input and output weights to prevent zero weights -Additional step for finding epsilon -Affected by outliers -Non-uniqueness of the solution -Not providing a full ranking of DMUs
Wang and Jiang	2012	-Selecting the best DMU in one step -Only the most efficient DMU can have an efficiency score of over 1, while the efficiency scores of others are all less than or equal to 1 -Corrected limitations of Foroughi (2011) -Lower bounds of weights are from slack-adjusted DEA models (Sueyoshi, 1999), thus no additional step for finding lower bounds	-Not providing a full ranking of DMUs -Cannot find the best-performing DMU in some cases

Sun et al.	2013	-Full ranking of DMUs	-Poor weight dispersion -Use of epsilon
Lam	2015	-Selecting the best DMU in one-step -Shows a better performance than the models by Foroughi (2011) and Wang & Jiang (2012)	-Additional step for finding epsilon -Not proposing a full ranking of DMUs -Cannot find the best DMU in some cases
Toloo	2015	-Lower bounds of weights are from slack-adjusted DEA models (Sueyoshi, 1999), thus no additional step for finding lower bounds	-Not providing a full ranking of DMUs for some cases -Cannot find the best DMU in some cases
Carillo and Jorge	2016	-Improved discriminatory characteristics	-Poor weight dispersion of inputs and outputs
Toloo and Salahi	2018	-Improved discriminating power	-Additional step for epsilon -The best-performing DMU could be incorrectly determined depending on the value of epsilon -Weights of inputs and outputs are very high for some cases
Karsak and Goker	2020	-Improved weight dispersion -Guarantees to identify best performing DMU	

Ozsoy et al.	2021	-Does not require epsilon -Computational advantages -Discriminating power high enough for full ranking	-Poor weight dispersion, the range of weight values in decimals is wide

4. METHODOLOGY

Data Envelopment Analysis (DEA) is a linear programming technique developed by Charnes et al. (1978) for measuring efficiency scores of homogenous decision-making units (DMUs) using multiple inputs and creating multiple outputs. DEA considers *n* decision-making units (DMUs) to be evaluated, where each DMU uses varying amounts of *m* different inputs to produce *s* different outputs. The relative efficiency (E_{j0}) of a DMU equals to the ratio of its total weighted output to its total weighted input. This ratio is the maximization objective of the mathematical program that is solved for the evaluated DMU. The model is completed with normalizing constraints that ensure the sum of weighted output to the sum of weighted input ratio of every DMU be less than or equal to unity.

$$\max E_{j0} = \frac{\sum_{r=1}^{s} u_r y_{r0}}{\sum_{i=1}^{m} v_i x_{i0}} \tag{4.1}$$

subject to

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1 \; ; \quad j = 1, \dots n,$$

$$u_r, v_i \geq 0; \quad r = 1, \dots s; \quad i = 1, \dots, m.$$

Here, y_{rj} is the amount of output *r* produced by DMU *j* and x_{ij} is the amount of input *i* used by DMU *j*, and u_r, v_i are the weights for output *r* and input *i*, respectively, which are determined in favour of the evaluated DMU in each computation. The subscript '*0*' is used for referencing the evaluated DMU. E_{j0} is calculated using the most favourable weights of outputs and inputs are (u_r and v_i) for maximizing the relative efficiency of the evaluated DMU.

The above model is an extended nonlinear programming formulation of an ordinary fractional programming problem. Thus, it can be rewritten as a linear programming model by using the transformation of linear fractional programming proposed by (Charnes & Cooper, 1962) and (Charnes & Cooper, 1973). The linear form of model (4.1):

$$\max E_{j0} = \sum_{r=1}^{s} u_r y_{rj0} \tag{4.2}$$

subject to

$$\sum_{i=1}^{m} v_i x_{ij0} = 1,$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, \forall j,$$

$$u_r, v_i \geq \varepsilon, \qquad \forall r, i.$$

where ε is a very small non-Archimedean positive number added to the model for preventing zero-weights.

As mentioned above, the conventional DEA model (CCR model) has several limitations such as weight flexibility and poor discriminating power. In order to deal with the limitations, researchers have focused on common-weight DEA-based models.

Karsak & Ahiska (2007) proposed a common-weight DEA-based model for identifying the best DMU in the context of multiple inputs and multiple outputs. The model has two stages. If the first stage results in identifying more than one efficient DMUs, the model proceeds with the second stage. The first stage of the model is as follows:

$$\min M \tag{4.3}$$

subject to

$$M - d_j \geq 0, \forall j,$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + d_j = 0, \forall j,$$

$$\sum_{r=1}^{s} u_r + \sum_{i=1}^{m} v_i = 1,$$

$$u_r, v_i, d_j \geq 0, \qquad \forall r, i, j.$$

where d_j represents the deviation of DMU j from the ideal efficiency score of 1, i.e. $d_j = 1 - E_j$, and M is the maximum deviation. Therefore, the model is a minimax efficiency model that aims to minimize the maximum deviation from ideal efficiency. If model (4.3) results in more than one efficient DMUs, then the model continues with the second stage as

$$\min M - k \textstyle\sum_{j \epsilon EF} d_j \tag{4.4}$$

subject to

$$M - d_j \geq 0, \forall j,$$

$$M - \sum_{j \epsilon EF} d_j \geq 0,$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + d_j = 0, \forall j,$$

$$\sum_{r=1}^{s} u_r + \sum_{i=1}^{m} v_i = 1,$$

$$u_r, v_i, d_j \geq 0, \qquad \forall r, i, j.$$

where $k \epsilon (0,1]$ is a discriminating parameter and its value is fixed by the analyst, and EF is the set of minimax efficient units determined in the first stage using formulation (4.3).

Another common weight approach aims to find the most efficient unit is the model proposed by Wang & Jiang (2012). The model is presented as follows:

$$min \sum_{i=1}^{m} v_i \left(\sum_{j=1}^{n} x_{ij} \right) - \sum_{r=1}^{s} u_r \left(\sum_{j=1}^{n} y_{rj} \right) \tag{4.5}$$

subject to

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq I_j, \qquad j = 1, \dots, n$$

$$\sum_{j=1}^{n} I_j = 1,$$

$$I_j \epsilon \{0,1\}, j = 1, \dots, n$$

$$u_r \geq \frac{1}{(m+s)max_j(y_{rj})}, \qquad r = 1, \dots, s$$

$$v_i \geq \frac{1}{(m+s)max_j(x_{ij})}, \qquad i = 1, \dots, m.$$

At the end, the model identifies one *DMU* as the most efficient one (DMU_k) and the binary variable I_j takes the value of 1 only for the most efficient unit (I_k =1).

Sun et al. (2013) introduced two mathematical programming models to assess common weight efficiency scores and to rank DMUs on a common basis. The first model is as given below:

$$min \sum_{j=1}^{n} d_j \tag{4.6}$$

subject to

$$\sum_{i=1}^{m} v_i x_{ij} - d_j = \sum_{r=1}^{s} u_r y_{rj}, \ \forall j,$$

$$\sum_{i=1}^{m} v_i x_{min} = 1,$$

$$\sum_{r=1}^{s} u_r y_{max} = 1,$$

$$u_r, v_i \geq \varepsilon, \forall r, i,$$

$$d_j \geq 0, \forall j.$$

where $x_{min} = min\{x_{ij} | j = 1, \dots, n\}, (i = 1, \dots, m)$ and $y_{max} = max\{y_{rj} | j = 1, \dots, n\}, (r = 1, \dots, s)$. As Sun et al. (2013) showed that model (4.6) could result in non-unique weights, they developed the following complementary model:

$$\max \sum_{i=1}^{m} v_i^2 + \sum_{r=1}^{s} u_r^2 \tag{4.7}$$

subject to

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} \right) = D^*,$$

$$\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} \geq 0, \forall j,$$

$$\sum_{i=1}^{m} v_i x_{min} = 1,$$

$$\sum_{r=1}^{s} u_r y_{max} = 1,$$

$$u_r, v_i \geq \varepsilon, \forall r, i.$$

where D^* represents the optimal value for the objective function of model (4.6). Sun et al. (2013) also suggested the following as an alternative for ranking common-weight efficiency scores:

$$min \sum_{j=1}^{n} \left(\sum_{i=1}^{m} v_i x_{max} - \sum_{i=1}^{m} v_i x_{ij} \right) + \sum_{j=1}^{n} \left(\sum_{r=1}^{s} u_r y_{rj} - \sum_{r=1}^{s} u_r y_{min} \right) \tag{4.8}$$

subject to

$$\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} \geq 0, \forall j,$$
$$\sum_{i=1}^{m} v_i x_{max} = 1,$$
$$\sum_{r=1}^{s} u_r y_{min} = 1,$$
$$u_r, v_i \geq \varepsilon, \forall r, i.$$

Models (4.7) and (4.8) enable full-ranking of DMUs; however, the weight dispersion of the models are very poor because most of the weights are equal to ε.

Another well-known technique for identifying the most efficient unit in a common-weight DEA context is proposed by Toloo (2015). The model is formulated as follows:

$$\min \; d_{max} \tag{4.9}$$

subject to

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + d_j - \beta_j = 0, \quad j = 1, \dots, n$$
$$d_{max} - d_j + \beta_j \geq 0, \quad j = 1, \dots, n$$
$$\sum_{j=1}^{n} d_j = n - 1,$$
$$d_j \epsilon \{0,1\}, \quad j = 1, \dots, n$$
$$\beta_j \leq 1, \quad j = 1, \dots, n$$
$$u_r \geq \frac{1}{(m+s)max_j(y_{rj})}, \quad r = 1, \dots, s$$
$$v_i \geq \frac{1}{(m+s)max_j(x_{ij})}, \quad i = 1, \dots, m.$$

In this model, $d_j - \beta_j$ is a deviation from the efficiency of DMU$_j$, and d_{max} is the maximum deviation, which is minimized. DMU$_k$ is the most efficient DMU if and only if $d_k^* - \beta_k^* = min\left\{d_j^* - \beta_j^*: j = 1,2, \dots, n\right\}$ or equivalently $d_k^* = 0$.

Carrillo & Jorge (2016) proposed a compromise programming approach for ranking alternatives. In their model, the concept of distance to an ideal is used as a means of selecting common-weights in favor of all DMUs simultaneously. The proposed model is elaborated by modeling the concept of distance to an ideal with Manhattan distance and Tchebychev distance. The model for Manhattan distance is given below.

$$min \sum_{j=1}^{n} \left(\sum_{i=1}^{m} v_i x_{ij} - m \right) + \left(M - \sum_{r=1}^{s} u_r y_{rj} \right) \tag{4.10}$$

subject to

$$m - \sum_{i=1}^{m} v_i x_{ij} \le 0,\ \forall j,$$
$$M - \sum_{r=1}^{s} u_r y_{rj} \ge 0, \forall j,$$
$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \forall j,$$
$$u_r, v_i \ge \varepsilon, \forall r, i,$$
$$m, M \ge 0.$$

where *m* is the minimum of weighted inputs and *M* is the maximum of weighted outputs. $(\sum_{i=1}^{m} v_i x_{ij} - m)$ represents the Manhattan distance between *m* and DMU *j* and $(M - \sum_{r=1}^{s} u_r y_{rj})$ represents the Manhattan distance between *M* and DMU *j*. This model measures the distance of DMU *j* to the ideal point (*m*,*M*). Carrillo & Jorge (2016) extended the proposed model by measuring the distance to the ideal point with Tchebychev distance.

$$min \sum_{j=1}^{n} \overline{D}_j \tag{4.11}$$

subject to

$$\overline{D}_j \quad - \sum_{i=1}^{m} v_i x_{ij} + m \geq 0, \forall j,$$

$$\overline{D}_j \quad - M + \sum_{r=1}^{s} u_r y_{rj} \geq 0, \forall j,$$

$$m - \sum_{i=1}^{m} v_i x_{ij} \leq 0, \ \forall j,$$

$$M - \sum_{r=1}^{s} u_r y_{rj} \geq 0, \forall j,$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, \forall j,$$

$$u_r, v_i \geq \varepsilon, \forall r, i,$$

$$\overline{D}_j \quad \geq 0, \forall j,$$

$$m, M \geq 0.$$

where $\overline{D}_j$ is the Tchebychev distance of DMU *j* to the ideal point. The objective of the model is to minimize the sum of distances to the ideal point.

Another recent common weight model is proposed by Toloo & Salahi (2018). In this model, there are two different steps. At the first step, the optimum value of the non-Archimedean ε^* is found and at the second step, the efficiency rankings are obtained. The continuous variable z_j is used for the linear transformation $(z_j = hI_j)$.
The model for finding ε^* is as follows:

$$\varepsilon^* = \max \varepsilon \tag{4.12}$$

subject to

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq MI_j - h + z_j, \quad j = 1, \dots, n$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \geq z_j - M(1 - I_j), \quad j = 1, \dots, n$$

$$\sum_{j=1}^{n} I_j = 1,$$

$$z_j \leq MI_j, \quad j = 1, \dots, n$$

$$z_j \leq h, \quad j = 1, \dots, n$$

$$h \leq z_j + M(1 - I_j), \quad j = 1, \dots, n$$

$$I_j \epsilon \{0,1\}, \quad j = 1, \dots, n$$

$$u_r \geq \varepsilon, \quad r = 1, \dots, s$$

$$v_i \geq \varepsilon, \quad i = 1, \dots, m$$

$$z_j \geq 0, \quad j = 1, \dots, n$$

$$h \geq \varepsilon$$

$$\varepsilon \geq 0.$$

After finding ε^*, the model continues as follows:

$$h^* = \max h \tag{4.13}$$

subject to

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq MI_j - h + z_j, \quad j = 1, \dots, n$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \geq z_j - M(1 - I_j), \quad j = 1, \dots, n$$

$$\sum_{j=1}^{n} I_j = 1,$$

$$z_j \leq MI_j, \quad j = 1, \dots, n$$

$$z_j \leq h, \quad j = 1, \dots, n$$

$$h \leq z_j + M(1 - I_j), \quad j = 1, \dots, n$$

$$I_j \epsilon \{0,1\}, \qquad j = 1, \dots, n$$
$$u_r \geq \varepsilon^*, \qquad r = 1, \dots, s$$
$$v_i \geq \varepsilon^*, \qquad i = 1, \dots, m$$
$$z_j \geq 0$$
$$h \geq 0.$$

where M is a very large positive number, h is the difference between the weighted sum of outputs and the weighted sum of inputs. z_j is a binary variable that takes the value of 1 just for the best DMU.

Karsak & Goker (2020) proposed a new common-weight DEA-based approach that ensures to discriminate the best performing DMU with an improved weight dispersion for inputs and outputs. Their approach is based on the idea of minimizing the maximum deviation from ideal efficiency. Moreover, their model is an improved version of the model proposed by Karsak and Ahiska (2007) that includes a non-Archimedean infinitesimal ϵ to the model and excludes the discriminating parameter *k*.

$$\min \theta \tag{4.14}$$

subject to

$$\theta - d_j \geq 0, \forall j,$$
$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + d_j = 0, \forall j,$$
$$\sum_{r=1}^{s} u_r + \sum_{i=1}^{m} v_i = 1,$$
$$u_r, v_i, \geq \varepsilon, \qquad \forall r, i,$$
$$d_j \geq 0, \forall j.$$

where θ is the maximum deviation from efficiency, and ε is a small positive scalar. If model (4.14) results in more than one minimax efficient unit, then the following formulation will be solved.

$$\min \theta \tag{4.15}$$

subject to

$$\theta - d_j \geq 0, \forall j,$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + d_j = 0, \forall j,$$

$$\sum_{r=1}^{s} u_r + \sum_{i=1}^{m} v_i = 1,$$

$$d_j + M z_j \geq \varepsilon, \qquad j \epsilon EF,$$

$$\sum_{j \epsilon EF} z_j = 1,$$

$$z_j \epsilon \{0,1\}, \qquad j \epsilon EF$$

$$u_r, v_i, \geq \varepsilon, \qquad \forall r, i,$$

$$d_j \geq 0, \forall j.$$

where is M a very large positive number and z_j is a binary variable, which equals to 1 only for the most efficient DMU.

Özsoy *et al.* (2021) suggested a new model based on formulations (4.13). The model is as follows:

$$h^* = \max h \tag{4.16}$$

subject to

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq M I_j - h + z_j, \qquad j = 1, \dots, n$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \geq z_j - M(1 - I_j), \qquad j = 1, \dots, n$$

$$\sum_{j=1}^{n} I_j = 1,$$

$$z_j \leq \; M I_j, \qquad j = 1, \dots, n$$

$$z_j \leq h, \qquad j = 1, \dots, n$$

$$h \leq \ z_j + M(1 - I_j), \qquad j = 1, \dots, n$$

$$I_j \epsilon \ \{0,1\}, \qquad j = 1, \dots, n$$

$$u_r \geq \frac{1}{(m+s)max_j(y_{rj})}, \qquad r = 1, \dots, s$$

$$v_i \geq \frac{1}{(m+s)max_j(x_{ij})}, \qquad i = 1, \dots, m$$

$$z_j \geq 0,$$

$$h \geq 0.$$

where M is a very large positive number, h is the difference between the weighted sum of outputs and the weighted sum of inputs. z_j is a binary variable that takes the value of 1 only for the best DMU. The proposed model (4.16) determines exactly one DMU as the most efficient one, other DMUs have efficiency scores strictly less than one, and it enables to rank all DMUs. Moreover, the model (4.16) does not use the non-Archimedean infinitesimal epsilon as the lower weight limit and consequently have some computational advantages such as fewer constraints.

5. EDUCATIONAL PERFORMANCE ASSESSMENT

5.1 Presentation of Data

This paper provides educational performance evaluation of OECD countries by taking into consideration some indicators from PISA 2018 database. The indicators are categorized as inputs and outputs in line with DEA context. DEA considers inputs as indicators to be minimized and outputs as indicators to be maximized.

For this study, the inputs and outputs are identified by taking into consideration the most frequently used indicators of PISA database in DEA applications (Aparicio *et al*., 2019). The education systems of countries are modeled as production processes that produce the success of students in PISA test by consuming different amounts of educational resources. The PISA test scores in reading, mathematics and science are the outputs and teachers/students ratio, learning time (in minutes per week) and the average number of computers available in schools for educational purposes are selected as the inputs of the study. Raw data for the inputs and outputs are shown in Table 5.1. Normalization for data is conducted via a linear normalization scheme. Normalization for input data is performed by using ${x_{ij}}/{x_i^*}$, where $x_i^* = max_j(x_{ij})$ for $\forall i,$ and output data is normalized as ${y_{rj}}/{y_r^*}$, where $y_r^* = max_j(y_{rj})$ for $\forall r$ (Karsak & Ahiska, 2007).

All the inputs and outputs are extracted from PISA 2018 database (OECD, 2018). Concerning the lack of data for some OECD countries, 29 countries are included in the analysis.

Table 5.1.1 : Input and Output Data for OECD Countries

OECD members	Input1	Input2	Input3	Output1	Output2	Output3
Australia	0.074627	1550	196	502.6317	491.36	502.9646
Austria	0.09177	1730	140	484.3926	498.9423	489.7804
Belgium	0.107527	1669	109	492.8644	508.0703	498.7731
Czech Republic	0.076555	1536	63	490.2188	499.4677	496.7913
Denmark	0.073529	1655	53	501.1299	509.3984	492.637
Estonia	0.080898	1563	56	523.017	523.4146	530.108
Finland	0.092001	1479.32	57	520.0787	507.3014	521.8846
France	0.086207	1646	183	492.6065	495.4076	492.9771
Greece	0.10492	1662	18	457.4144	451.3703	451.6327
Hungary	0.093966	1586	65	475.9867	481.0826	480.9117
Iceland	0.09423	1637	68	473.9743	495.1874	475.0241
Ireland	0.078171	1730	77	518.0785	499.6329	496.1136
Israel	0.085977	1753	62	470.4152	463.0345	462.1966
Italy	0.107224	1742	90	476.2847	486.5899	468.0117
Japan	0.083785	1682.94	108	503.856	526.9733	529.1354
Korea	0.074331	1776.4	92	514.0523	525.933	519.0073
Latvia	0.09116	1556	42	478.6987	496.1263	487.2506
Lithuania	0.095214	1510	64	475.8735	481.1912	482.067
Luxembourg	0.106603	1646	213	469.9854	483.4215	476.7694
Netherlands	0.057406	1632	199	484.7837	519.231	503.3838
Poland	0.123712	1697	33	511.8557	515.6479	511.0356
Portugal	0.096169	1759	69	491.8008	492.4874	491.6773
Slovak Republic	0.077637	1517	56	457.984	486.1649	464.0476
Slovenia	0.113344	1686	71	495.3456	508.8975	507.0065
Sweden	0.083958	1631	106	505.7852	502.3877	499.4447
Switzerland	0.077259	1565	83	483.9294	515.3147	495.2763
Turkey	0.074335	1571.39	38	465.6317	453.5078	468.2996
United Kingdom	0.065359	1613	260	503.9281	501.7699	504.6675
United States	0.05814	1822	475	505.3528	478.2447	502.38

5.2. Results

This section shows the implementation of various common-weight DEA-based approaches using PISA data and provides a comparative evaluation for different outcomes obtained by the implemented models. CPLEX solver in GAMS software is used for performing computations of the models. For this study, the scalar *M* equals to 10^6 and the epsilon value is taken to be 10^{-6} except the model proposed by Toloo & Salahi (2018) where the epsilon value is calculated as 1.2009E+5 with a mathematical programming model.

Overall ranking results and efficiency scores are shown in Table 5.2.2. The CCR model yields ten efficient countries that are DMU_5, DMU_6, DMU_7, DMU_9, DMU_{17}, DMU_{20}, DMU_{21}, DMU_{27}, DMU_{28}, DMU_{29}, while other DMUs are inefficient. In order to allow further discrimination among DMUs, models proposed by Karsak & Ahiska (2007), model by Wang & Jiang (2012), models by Sun *et al.* (2013), model by Toloo (2015), model by Carillo & Jorge (2016), models by Toloo & Salahi (2018), models by Karsak & Goker (2020), and model by Özsoy *et al*. (2021) are applied to PISA data. These models were detailed in Section 4.

As seen in Table 5.2.2, all common-weight models other than the model by Wang & Jiang (2012) achieved full ranking of DMUs. Estonia (DMU_6) is identified as the most efficient country according to model (4.4), model (4.5), model (4.9), and model (4.11). On the other hand, Finland (DMU_7) is identified as the best performing country by model (4.7), model (4.8), and model (4.15) while Estonia has taken the second place in these models' results. Alternatively, Netherlands (DMU_{20}) is determined as the most efficient country by model (4.13) and model (4.16).

Estonia and Finland are identified as minimax efficient countries according to models (4.3) and (4.14). At the second step of model (4.3), Estonia is the most efficient country for discriminating parameter *k* equals to 0.47. However, Finland is the most efficient unit with regard to the second step of model (4.14). Model (4.5) proposed by Wang & Jiang

(2012) assigns efficiency score of one to Czech Republic and Finland. Alternatively, model (4.9) also calculates efficiency score of one for Czech Republic.

In Table 5.2.1 the weight dispersion for inputs and outputs obtained by various common-weight DEA-based models is given. In model (4.7), v_1, u_1 and u_3 equal to epsilon (ε) and v_3 is very close to epsilon with a value of 2.3991E-5, thus this model results in poor weight dispersion by taking into account just one input and one output for efficiency evaluation. In a same way, model (4.8) assigns the value of epsilon to v_1, v_3, u_2 and u_3. Hence, model (4.8) considers just one input and one output. Models (4.5) and (4.9) achieve very similar weights as they used the weight constraints from (Sueyoshi, 1999) for avoiding zero-weights. Model (4.11) yields poor weight dispersion because beside the value of v_2, which equals to 2.60E-06; all the weights equal to 0.000001, which is the epsilon value. In model (4.13), u_1 and u_3 equal to epsilon (ε), therefore this model considers just one output ($u_2 > \varepsilon$). The mentioned unrealistic weighting schemes that consider one input and one output disregard the robustness of the related models.

Models (4.4) and (4.15) yield enhanced weight dispersion by considering all the inputs and two outputs for efficiency calculations. On the other hand, model (4.16) takes into account all the inputs and outputs with non-zero weights but the magnitude of weights varies considerably. For example, u_1 and u_3 equal to 0.167 while u_2 equals to 2.39E+06. The importance of the second output is relatively very high in efficiency calculations.

Table 5.2.1: Weight Dispersion for Inputs and Outputs

Weights	Karsak & Ahiska (2007)	Wang & Jiang (2012)	Sun *et al*. (2013) model 1	Sun *et al*. (2013) model 2	Toloo (2015)	Carrillo & Jorge (2016) Tchebychev	Toloo & Salahi (2018)	Karsak & Goker (2020)	Özsoy *et al*. (2021)
v1	0.18	0.261	0.000001	0.000001	0.271	0.000001	2.2155E+06	0.103	2.19E+06
v2	0.339	0.34	1.232	1	0.332	2.60E-06	1.2027E+06	0.427	1.20E+06
v3	0.063	0.167	2.3991E-05	0.000001	0.167	0.000001	3.3492E+05	0.03	3.17E+05
u1	0.008	0.167	0.000001	1.143	0.167	0.000001	1.2009E+05	0.133	0.167
u2	0.41	0.167	1	0.000001	0.167	0.000001	2.1723E+06	0.307	2.39E+06
u3	0	0.167	0.000001	0.000001	0.167	0.000001	1.2009E+05	0.000001	0.167

Table 5.2.2: Efficiency Scores and Ranking Results

		CCR efficiency	ranking	Karsak & Ahiska (2007)_step 1	ranking	Karsak & Ahiska (2007) step_2 **k=0.47**	ranking	Wang & Jiang (2012)	ranking	Sun *et al.* (2013)_model 1	ranking	Sun *et al.* (2013)_model 2	ranking	Toloo (2015)	ranking	Carrillo & Jorge(2016) (Tchebychev)	ranking	Toloo & Salahi (2018)	ranking	Karsak & Goker (2020)_step 1	ranking	Karsak & Goker (2020)_step 2	ranking	Özsoy et al. (2021)	ranking
Country1	Australia	0.9902	12	0.94363	6	0.923099	11	0.918766	14	0.8899	9	1.291686	3	0.919167	14	0.88135	15	0.902738	9	0.9436	6	0.9436	6	0.898482	9
Country2	Austria	0.85868	25	0.84154	23	0.835427	21	0.824011	24	0.8096	24	1.115296	24	0.823318	22	0.79864	24	0.790164	21	0.8415	23	0.8415	23	0.789484	21
Country3	Belgium	0.88769	23	0.86228	20	0.837338	20	0.82294	25	0.8546	14	1.176278	19	0.81997	25	0.81886	22	0.747553	24	0.8623	20	0.8623	20	0.74689	24
Country4	Czech Republic	0.98329	14	0.96976	3	0.977562	4	1	2	0.9128	4	1.271270	4	1	2	0.95978	3	0.940201	6	0.9698	3	0.9698	3	0.937437	6
Counry5	Denmark	1	1	0.93564	8	0.958299	6	0.985453	4	0.864	13	1.206123	10	0.987024	4	0.93132	5	0.950914	3	0.9356	8	0.9356	8	0.948933	3
Country6	Estonia	1	1	1	1	1	1	1.036943	1	0.9401	2	1.332895	2	1.036389	1	1	1	0.952304	2	1	1	1	2	0.947726	5
Country7	Finland	1	1	1	1	0.96703	5	1	2	0.9627	1	1.400384	1	0.996813	3	0.98977	2	0.874111	12	1	1	1	1	0.868435	12
Country8	France	0.89749	20	0.88027	18	0.862789	19	0.846896	19	0.8449	19	1.192090	16	0.846349	19	0.8203	21	0.821553	17	0.8803	18	0.8803	18	0.819808	17
Country9	Greece	1	1	0.79186	29	0.771905	29	0.800308	26	0.7624	27	1.096271	27	0.797499	26	0.79366	26	0.691933	29	0.7919	29	0.7919	29	0.688466	29
Country10	Hungary	0.89043	22	0.88201	17	0.866992	18	0.879917	17	0.8515	16	1.195448	13	0.877745	17	0.86488	16	0.793139	20	0.882	17	0.882	17	0.790515	20
Country11	Iceland	0.89138	21	0.87513	19	0.871544	17	0.865099	18	0.8492	17	1.153308	21	0.863326	18	0.84686	18	0.801738	18	0.8751	19	0.8751	19	0.801958	18
Country12	Ireland	0.96453	15	0.88751	15	0.890515	15	0.930415	11	0.8107	22	1.192857	14	0.931665	11	0.88189	14	0.882434	11	0.8875	15	0.8875	15	0.8774	11
Country13	Israel	0.83031	28	0.79895	28	0.800371	25	0.832674	21	0.7415	28	1.068903	29	0.832861	21	0.79727	25	0.77409	22	0.799	28	0.799	28	0.770233	22
Country14	Italy	0.81571	29	0.80231	27	0.784749	26	0.776705	27	0.7841	26	1.089074	28	0.77438	27	0.76773	27	0.708851	26	0.8023	27	0.8023	27	0.708391	26
Country15	Japan	0.94064	17	0.92186	11	0.930422	8	0.933324	10	0.879	10	1.192554	15	0.933338	10	0.89659	9	0.895951	10	0.9219	11	0.9219	11	0.894471	10
Country16	Korea	0.98491	13	0.90117	14	0.925919	10	0.947329	8	0.8311	20	1.152672	22	0.949551	8	0.88989	12	0.935768	7	0.9012	14	0.9012	14	0.933539	7
Country17	Latvia	1	1	0.92487	10	0.9202	12	0.929866	12	0.8951	6	1.225439	9	0.927689	12	0.91226	7	0.842406	16	0.9249	10	0.9249	10	0.841142	16
Country18	Lithuania	0.92926	18	0.91605	13	0.892384	14	0.901378	15	0.8946	7	1.255317	5	0.898334	16	0.89398	10	0.801623	19	0.9161	13	0.9161	13	0.799001	19
Country19	Luxembourg	0.85643	26	0.82016	25	0.783622	27	0.745669	29	0.8245	21	1.137348	23	0.743089	29	0.74272	28	0.702112	28	0.8202	25	0.8202	25	0.702039	27
Country20	Netherlands	1	1	0.95921	5	0.996089	2	0.962322	6	0.8931	8	1.183224	17	0.966325	6	0.89184	11	1.05348	1	0.9592	5	0.9592	5	1.0557380	1
Country21	Poland	1	1	0.8546	21	0.818235	24	0.826093	23	0.853	15	1.201447	12	0.821414	24	0.8375	19	0.702284	27	0.8546	21	0.8546	21	0.700114	28
Country22	Portugal	0.83049	27	0.82773	24	0.820801	23	0.840821	20	0.786	25	1.113685	25	0.839718	20	0.81714	23	0.768918	23	0.8277	24	0.8277	24	0.765945	23
Country23	Slovak Republic	0.96128	16	0.94138	7	0.957582	7	0.954886	7	0.8996	5	1.202552	11	0.95453	7	0.91957	6	0.910812	8	0.9414	7	0.9414	7	0.911715	8
Country24	Slovenia	0.88017	24	0.85337	22	0.827339	22	0.826232	22	0.8473	18	1.170280	20	0.822658	23	0.82722	20	0.728638	25	0.8534	22	0.8534	22	0.72709	25
Country25	Sweden	0.92828	19	0.91618	12	0.907412	13	0.91958	13	0.8647	12	1.235239	7	0.919175	13	0.88727	13	0.866408	14	0.9162	12	0.9162	12	0.863259	13
Country26	Switzerland	0.9904	11	0.96799	4	0.985743	3	0.976489	5	0.9243	3	1.231705	8	0.976617	5	0.93656	4	0.948453	5	0.968	4	0.968	4	0.94949	2
Country27	Turkey	1	1	0.88269	16	0.888282	16	0.94718	9	0.8102	23	1.180317	18	0.947879	9	0.90204	8	0.869407	13	0.8827	16	0.8827	16	0.862873	14
Country28	United Kingdom	1	1	0.93443	9	0.927788	9	0.901332	16	0.8732	11	1.244438	6	0.90353	15	0.84981	17	0.950344	4	0.9344	9	0.9344	9	0.948197	4
Country29	United States	1	1	0.80329	26	0.782071	28	0.746997	28	0.7368	29	1.104803	26	0.750338	28	0.69381	29	0.853606	15	0.8033	26	0.8033	26	0.849445	15

To conclude whether there is a similarity between the sets of rankings of these common-weight DEA-based approaches, Spearman's rank correlation is conducted between pairs of different models. The null hypothesis (H_0) and alternative hypothesis H_a are defined as follows:

H_0: There is no correlation between the sets of rankings obtained by the compared models.
H_a: There is a positive correlation between the sets of rankings obtained by the compared models.

The results of these tests are reported in Table 5.2.3 in matrix format. For n=29 and α=0.05; the benchmark value $r_{s,\alpha}$ is 0.368 (Ramsey, 1989). Thus, the null hypothesis is rejected, the similarity between rankings of the models is statistically significant. If r_s is close to +1, the correlation between the rankings becomes more significant as $r_s = 1$ represents the perfect correlation situation.

Table 5.2.3: Spearman Rank Correlation Coefficients

Spearman Rank Coef.	Karsak & Ahiska (2007)	Wang & Jiang (2012)	Sun *et al.* (2013) model1	Sun *et al.* (2013) model2	Toloo (2015)	Carrillo & Jorge (2016)	Toloo & Salahi (2018)	Karsak & Goker (2020)	Özsoy *et al.* (2021)
Karsak & Ahiska (2007)	1	0.9408	0.8517	0.7655	0.9453	0.9143	0.9015	0.969	0.8985
Wang & Jiang (2012)		1	0.7403	0.7068	0.9979	0.9585	0.8553	0.9085	0.8445
Sun et al. (2013) model1			1	0.8872	0.73	0.8355	0.6094	0.9099	0.6103
Sun et al. (2013) model2				1	0.6975	0.7961	0.602	0.8778	0.5931
Toloo (2015)					1	0.9502	0.869	0.9084	0.8586
Carrillo & Jorge (2016)						1	0.7547	0.9148	0.7433
Toloo & Salahi (2018)							1	0.8517	0.9946
Karsak & Goker (2020)								1	0.8488
Özsoy et al. (2021)									1

The rank correlation coefficient between the models by Karsak & Ahiska (2007) and Karsak & Goker (2020) is close to +1, thus they result in very similar rankings. As these models are minimax efficiency models, it is obvious to achieve this similarity. Moreover, the models by Wang & Jiang (2012) and Toloo (2015) have similar rankings as Toloo (2015) model is based on Wang & Jiang (2012)'s model. Alternatively, these two models give similar results with Carrillo & Jorge (2016) model with rank correlation coefficients 0.9585 and 0.9502, respectively. Another strong correlation is between models Toloo & Salahi (2018) and Özsoy *et al.* (2021) because models by Özsoy *et al.* (2021) are inspired by models developed by Toloo & Salahi (2018).

5.3. Novel Approach for Comparing Results

In order to identify the best-performing DMU, a new mathematical programming model is proposed in this thesis. The model consists of solving *(n-1)* mixed integer linear programming models, where *n* is the number of DMUs, to determine the full ranking of alternatives. In the solution process, the model starts by evaluating *n* DMUs and at each iteration, one DMU is eliminated. The eliminated unit is the best DMU of the related iteration. The developed model is as the following:

$$\min \theta \tag{5.3.1}$$

subject to

$$\frac{U_r - f_r}{U_r - L_r} \leq \theta, \qquad for \;\; r = 1,2, \dots s$$

$$\frac{f_i - L_i}{U_i - L_i} \leq \theta, \qquad for \;\; i = 1,2, \dots m$$

$$\sum_{j=1}^{n} a_j = 1,$$

$$\theta \geq 0,$$

$$a_j \in \{0,1\}, \qquad j = 1,2, \dots, n.$$

where θ is the maximum deviation from the ideal situation. f_r is the function representing the *r* th output and f_i is the function representing the *i* th input. f_r equals to $\sum_j a_j y_{rj}$ and f_i equals to $\sum_j a_j x_{ij}$. The binary variable a_j equals to 1 just for the best alternative. U_r is the maximum amount of output and can be considered as the upper bound of output *r* and equals to $y^*_{rmax} = max_j(y_{rj})$. L_r is the lower bound of output *r* and equals to $y^*_{rmin} = min_j(y_{rj})$. In a same way, U_i and L_i are the upper and lower bounds of input *i*, respectively. U_i equals to $x^*_{imax} = max_j(x_{ij})$ and L_i equals to $x^*_{rmin} = min_j(x_{ij})$. At each iteration, the upper and lower bounds are updated.

This model is applied to PISA dataset and the ranking results are presented and compared with other models in Table 5.3.1.

Table 5.3.1. Rankings Obtained with respect to Different Models

		CCR efficiency	Karsak & Ahiska (2007)	Wang & Jiang (2012)	Sun *et al.* (2013) model 1	Sun *et al.* (2013) model 2	Toloo (2015)	Carrillo & Jorge (2016)	Toloo & Salahi (2018)	Karsak & Goker (2020)	Özsoy *et al.* (2021)	NEW MODEL
Country 1	Australia	12	11	14	9	3	14	15	9	6	9	2
Country 2	Austria	25	21	24	24	24	22	24	21	23	21	16
Country 3	Belgium	23	20	25	14	19	25	22	24	20	24	19
Country 4	Czech Republic	14	4	2	4	4	2	3	6	3	6	4
Country 5	Denmark	1	6	4	13	10	4	5	3	8	3	6
Country 6	Estonia	1	1	1	2	2	1	1	2	2	5	1
Country 7	Finland	1	5	2	1	1	3	2	12	1	12	7
Country 8	France	20	19	19	19	16	19	21	17	18	17	5
Country 9	Greece	1	29	26	27	27	26	26	29	29	29	27
Country10	Hungary	22	18	17	16	13	17	16	20	17	20	13
Country11	Iceland	21	17	18	17	21	18	18	18	19	18	14
Country12	Ireland	15	15	11	22	14	11	14	11	15	11	17
Country13	Israel	28	25	21	28	29	21	25	22	28	22	21
Country14	Italy	29	26	27	26	28	27	27	26	27	26	18
Country15	Japan	17	8	10	10	15	10	9	10	11	10	11
Country16	Korea	13	10	8	20	22	8	12	7	14	7	24
Country17	Latvia	1	12	12	6	9	12	7	16	10	16	12
Country18	Lithuania	18	14	15	7	5	16	10	19	13	19	15
Country19	Luxembourg	26	27	29	21	23	29	28	28	25	27	20
Country20	Netherlands	1	2	6	8	17	6	11	1	5	1	10
Country21	Poland	1	24	23	15	12	24	19	27	21	28	28

Country22	Portugal	27	23	20	25	25	20	23	23	24	23	22
Country23	Slovak Republic	16	7	7	5	11	7	6	8	7	8	26
Country24	Slovenia	24	22	22	18	20	23	20	25	22	25	23
Country25	Sweden	19	13	13	12	7	13	13	14	12	13	3
Country26	Switzerland	11	3	5	3	8	5	4	5	4	2	9
Country27	Turkey	1	16	9	23	18	9	8	13	16	14	25
Country28	United Kingdom	1	9	16	11	6	15	17	4	9	4	8
Country29	United States	1	28	28	29	26	28	29	15	26	15	29

According to the novel approach, Estonia is identified as the best alternative in terms of education performance. In the majority of the models, Finland and Estonia shared the first two rankings, however in the new model Australia takes the second rank. Moreover, Australia is followed by Sweden and Czech Republic in the new model. The rankings are compared with Spearman rank correlations as in Table 5.3.2.

Table 5.3.2. Spearman Rank Correlations of the Novel Approach

Spearman Rank Correlation	New model
Karsak & Ahiska (2007)	0.668472906
Wang & Jiang (2012)	0.557019704
Sun et al. (2013) model 1	0.637438424
Sun et al. (2013) model 2	0.698029557
Toloo (2015)	0.57044335
Carrillo & Jorge (2016)	0.530541872
Toloo & Salahi (2018)	0.570935961
Karsak & Goker (2020)	0.698029557
Özsoy et al. (2021)	0.573891626

As seen in Table 5.3.2, the new rankings are very similar with rankings obtained by the second model of Sun *et al.* (2013) and the model proposed by Karsak & Goker (2020) with a correlation coefficient of 0.70. Another strong correlation is calculated between the rankings obtained by the new model and Karsak & Ahiska (2007)'s model with a coefficient of 0.67. In the models of Karsak &Ahiska (2007) and Karsak & Goker (2020), the objective is to minimize the maximum deviation from efficiency. Therefore, it can be considered that the objective of the new model is similar to these models. Finally, it is obvious to say that the similarity between the rankings is statistically significant.

6. DISCUSSION & CONCLUDING REMARKS

Governments are looking forward to enhance their educational policies for managing educational resources efficiently and for establishing successful education systems. In order to measure these efforts, the PISA proposed by OECD provides a valuable source of data. PISA measures the success of 15-year-old students in main subjects of reading, mathematics and science every three years and provides various indicators that influence the students' success. Thus, the latest edition of PISA 2018 database assesses the current situation of education systems of OECD countries and enables to derive policies for improving the existing systems. In this context, PISA database can be a valuable source for measuring how efficient are the education systems of OECD countries.

This thesis compares educational performance of OECD countries by using different indicators from PISA 2018 database. The performance assessment is conducted by various common-weight DEA-based approaches. As conventional DEA models have several limitations such as poor weight dispersion and insufficient discriminatory characteristics, the use of common-weight DEA-based models is practical in terms of efficiency evaluation and ranking.

The models proposed by Karsak & Ahiska (2007), Wang & Jiang (2012), Sun *et al.* (2013), Toloo (2015), Carillo & Jorge (2016), Toloo & Salahi (2018), Karsak & Goker (2020) and Özsoy *et al.* (2021) that are selected from literature on common-weight DEA-based methods by considering relative advantages and limitations are implemented. The outputs are selected as reading, mathematics and science scores in PISA tests while the inputs are number of teachers per number of students, total learning time in minutes per week and average number of computers available in schools for education purposes.

The results obtained by different models are compared in terms of efficiency scores, rankings and weight dispersions for inputs and outputs. In majority of the models, Estonia and Finland are selected as the most efficient countries. The similarities between the sets of rankings obtained by different models are tested with Spearman rank correlation. The strongest similarity is observed between the rank-orders obtained by Wang & Jiang (2012) and Toloo (2015); and Toloo & Salahi (2018) and Özsoy *et al.* (2021). The problem of obtaining a full rank-order is observed for the CCR model as it identifies ten countries as efficient. Moreover, the model of Wang & Jiang (2012) cannot fully rank the DMUs by yielding the same efficiency score of one for two countries. The weight dispersions are enhanced in models proposed by Karsak & Ahiska (2007) and Karsak & Goker (2020) as these models consider all the outputs and two inputs for efficiency calculations. Models proposed by Sun et al. (2013), Carrillo & Jorge (2016) and Toloo & Salahi (2018) yield poor weight dispersion by assigning a negligible value to more than one input and/or output.

To the best of our knowledge, this thesis will be the first one to implement common-weight DEA-based models in the context of educational performance assessment using PISA data. Until now, efficiency assessments in education are generally conducted with conventional DEA models, thus a full rank-order could not be achieved among DMUs. Therefore, this study provides a more comprehensive and practical assessment among OECD countries by providing a rank-order for all the countries in most of the applied models. This assessment can be good reference for policy-makers to observe the current status of their countries in terms of education efficiency. Moreover, as the PISA 2018 is the latest edition of the PISA cycle, the performance assessment provides the latest analysis.

Although the PISA database is very rich in terms of educational indicators and indexes, it contains only exact data. However, PISA also publishes the answers of students to different questions used in PISA tests and in extra questionnaires. Therefore, the future research may focus on proposing a methodology to deal with these answers in a way to incorporate imprecision into the decision framework. This will enable to include qualitative data provided as linguistic variables into the education performance assessment.

REFERENCES

Afonso, A., & St. Aubyn, M. (2006). Cross-country efficiency of secondary education provision : A semi-parametric analysis with non-discretionary inputs. *Economic Modelling*, *23*, 476–491. https://doi.org/10.1016/j.econmod.2006.02.003

Agasisti, T. (2014). The Efficiency of Public Spending on Education : an empirical comparison of EU countries. *European Journal of Education*, *49*(4). https://doi.org/10.1111/ejed.12069

Agasisti, T., & Zoido, P. (2019). The efficiency of schools in developing countries , analysed through PISA 2012 data. *Socio-Economic Planning Sciences*, *68*.

Allen, R., Athanassopoulos, A., Dyson, R. G., & Thanassoulis, E. (1997). Weights restrictions and value judgements in Data Envelopment Analysis: Evolution, development and future directions. *Annals of Operations Research*, *73*, 13–34. https://doi.org/10.1023/a:1018968909638

Andersen, P., & Petersen, N. C. (1993). A Procedure for Ranking Efficient Units in Data Envelopment Analysis. *Management Science*, *39*(10), 1261–1264. https://doi.org/10.1287/mnsc.39.10.1261

Aparicio, J., Cordero, J. M., Gonzalez, M., & Lopez-espin, J. J. (2018). Using non-radial DEA to assess school efficiency in a cross-country perspective : An empirical analysis of OECD countries R. *Omega*, *79*, 9–20. https://doi.org/10.1016/j.omega.2017.07.004

Aparicio, J., Cordero, J. M., & Ortiz, L. (2019). Measuring efficiency in education : The influence of imprecision and variability in data on DEA estimates. *Socio-Economic Planning Sciences*, *68*.

Aparicio, J., Cordero, J. M., & Pastor, J. T. (2017). The determination of the least distance to the strongly efficient frontier in Data Envelopment Analysis oriented models: Modelling and computational aspects. *Omega*, *71*, 1–10. https://doi.org/10.1016/j.omega.2016.09.008

Aparicio, J., Perelman, S., & Santín, D. (2020). Comparing the evolution of productivity and performance gaps in education systems through DEA: an application to Latin American countries. *Operational Research*. https://doi.org/10.1007/s12351-020-00578-2

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

Ben Yahia, F., Essid, H., & Rebai, S. (2018). Do dropout and environmental factors matter? A directional distance function assessment of tunisian education efficiency. *International Journal of Educational Development*, *60*(April 2017), 120–127. https://doi.org/10.1016/j.ijedudev.2017.11.004

Bogetoft, P., Heinesen, E., & Tranæs, T. (2015). The efficiency of educational production: A comparison of the Nordic countries with other OECD countries. *Economic Modelling*, *50*, 310–321. https://doi.org/10.1016/j.econmod.2015.06.025

Carrillo, M., & Jorge, J. M. (2016). A multiobjective DEA approach to ranking alternatives. *Expert Systems with Applications*, *50*, 130–139. https://doi.org/10.1016/j.eswa.2015.12.022

Charnes, A., & Cooper, W. W. (1962). Programming with linear fractional functionals. *Naval Research Logistics Quarterly*, *9*(3,4), 181–185.

Charnes, A., & Cooper, W. W. (1973). An explicit general solution in linear fractional programming. *Naval Research Logistics Quarterly*, *20*(3), 449–467. https://doi.org/10.1002/nav.3800200308

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

Coco, G., & Lagravinese, R. (2014). Cronyism and education performance. *Economic Modelling*, *38*, 443–450.

Coco, G., Lagravinese, R., & Resce, G. (2020). Beyond the weights: a multicriteria approach to evaluate inequality in education. *Journal of Economic Inequality*, *18*(4), 469–489. https://doi.org/10.1007/s10888-020-09449-4

Delfin-Ortega, O. V., & Navarro-Chávez, C. L. (2020). The efficiency of higher education in Mexico, 2008-2016: A dynamic-network DEA model. *Perfiles Latinoamericanos*, *28*(56). https://doi.org/10.18504/pl2856-011-2020

Deutsch, J., Dumas, A., & Silber, J. (2019). School externalities and scholastic performance: an efficiency analysis. *International Journal of Manpower*, *40*(1), 102–119. https://doi.org/10.1108/IJM-09-2017-0220

Doyle, J., & Green, R. (1994). Efficiency and cross-efficiency in DEA derivations, meanings and uses. *Journal of the Operational Research Society*, *45*(5), 567–578. https://doi.org/10.1057/jors.1994.84

Foladi, S., Solimanpur, M., & Jahangoshai Rezaee, M. (2020). Inverse Dynamic Data Envelopment Analysis for Evaluating Faculties of University with Quasi-Fixed Inputs. *Social Indicators Research*, *148*(1), 323–347. https://doi.org/10.1007/s11205-019-02196-8

Foroughi, A. A. (2011). A new mixed integer linear model for selecting the best decision making units in data envelopment analysis. *Computers and Industrial Engineering*, *60*(4), 550–554. https://doi.org/10.1016/j.cie.2010.12.012

Karsak, E. E., & Ahiska, S. S. (2005). Practical common weight multi-criteria decision-making approach with an improved discriminating power for technology selection. *International Journal of Production Research*, *43*(8), 1537–1554. https://doi.org/10.1080/13528160412331326478

Karsak, E. Ertugrul, & Ahiska, S. S. (2007). A common-weight MCDM framework for decision problems with multiple inputs and outputs. *Lecture Notes in Computer Science*, *1*, 779–790. https://doi.org/10.1007/978-3-540-74472-6_64

Karsak, E. Ertugrul, & Goker, N. (2020). Improved common weight DEA-based decision approach for economic and financial performance assessment. *Technological and Economic Development of Economy*, *26*(2), 430–448. https://doi.org/10.3846/tede.2020.11870

Lagravinese, R., Liberati, P., & Resce, G. (2020). The impact of economic, social and cultural conditions on educational attainments. *Journal of Policy Modeling*, *42*(1), 112–132. https://doi.org/10.1016/j.jpolmod.2019.03.007

Lam, K. F. (2015). In the determination of the most efficient decision making unit in data envelopment analysis. *Computers and Industrial Engineering*, *79*, 76–84. https://doi.org/10.1016/j.cie.2014.10.027

Mancebón, M. J., Calero, J., Choi, Á., & Ximénez-De-Embún, D. P. (2012). The efficiency of public and publicly subsidized high schools in Spain: Evidence from

PISA-2006. *Journal of the Operational Research Society*, *63*(11), 1516–1533. https://doi.org/10.1057/jors.2011.156

Margaritis, S. G., Tsamadias, C. P., & Argyropoulos, E. E. (2021). Investigating the Relative Efficiency and Productivity Change of Upper Secondary Schools: the Case of Schools in the Region of Central Greece. *Journal of the Knowledge Economy*. https://doi.org/10.1007/s13132-020-00698-2

Marti Selva, M. L., & Medina, R. P. (2018). Comparison of educational efficiency in Europe and Asia : TIMMS 2015. *Revista de Educacion*, (380). https://doi.org/10.4438/1988-592X-RE-2017-380-372

Mou, H., Atkinson, M. M., & Marshall, J. (2019). Budgeting for efficiency? A case study of the public K-12 education systems of Canada. *Applied Economics*, *51*(34), 3740–3757. https://doi.org/10.1080/00036846.2019.1584380

OECD. (2017). *PISA 2015 Technical Report*. Paris.

OECD. (2018). PISA 2018 Database. Retrieved from https://pisadataexplorer.oecd.org/ide/idepisa/dataset.aspx

OECD. (2019a). *Education at a Glance 2019: OECD Indicators*. Paris. https://doi.org/https://doi.org/10.1787/f8d7880d-en

OECD. (2019b). *PISA 2018 Assessment and Analytical Framework*. *PISA OECD Publishing*. Par. https://doi.org/https://doi.org/10.1787/b25efab8-en

OECD. (2019c). *PISA 2018 Results (Volume I) : What Students Know and Can Do*. *OECD Publishing*. Paris. https://doi.org/https://doi.org/10.1787/5f07c754-en

Özsoy, V. S., Örkcü, H. H., & Örkcü, M. (2021). A simplistic approach without epsilon to choose the most efficient unit in data envelopment analysis. *Expert Systems with Applications*, *168*(December 2020). https://doi.org/10.1016/j.eswa.2020.114472

Ramsey, P. (1989). Critical Values for Spearman's Rank Order Correlation. *Journal of Educational Statistics*, *14*, 245–253.

Rebai, S., Ben Yahia, F., & Essid, H. (2020). A graphically based machine learning approach to predict secondary schools performance in Tunisia. *Socio-Economic Planning Sciences*, *70*(July 2019). https://doi.org/10.1016/j.seps.2019.06.009

Santín, D., & Sicilia, G. (2015). Measuring the efficiency of public schools in Uruguay: Main drivers and policy implications. *Latin American Economic Review*, *24*(1). https://doi.org/10.1007/s40503-015-0019-5

Segovia-Gonzalez, M. M., Dominguez, C., & Contreras, I. (2020). An assessment of the efficiency of Spanish schools: evaluating the influence of the geographical, managerial, and socioeconomic features. *International Transactions in Operational Research*, *27*(4), 1845–1868. https://doi.org/10.1111/itor.12711

Shamohammadi, M., & Oh, D. hyun. (2019). Measuring the efficiency changes of private universities of Korea: A two-stage network data envelopment analysis. *Technological Forecasting and Social Change*, *148*(September). https://doi.org/10.1016/j.techfore.2019.119730

Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage , semi-parametric models of production processes. *Journal of Econometrics*, *136*, 31–64. https://doi.org/10.1016/j.jeconom.2005.07.009

Su, W., Wang, D., Xu, L., Zeng, S., & Zhang, C. (2020). A Nonradial Super Efficiency DEA Framework Using a MCDM to Measure the Research Efficiency of Disciplines at Chinese Universities. *IEEE Access*, *8*, 86388–86399. https://doi.org/10.1109/ACCESS.2020.2993108

Sueyoshi, T. (1999). DEA non-parametric ranking test and index measurement: Slack-adjusted DEA and an application to Japanese agriculture cooperatives. *Omega*, *27*(3), 315–326. https://doi.org/10.1016/S0305-0483(98)00057-7

Sun, J., Wu, J., & Guo, D. (2013). Performance ranking of units considering ideal and anti-ideal DMU with common weights. *Applied Mathematical Modelling*, *37*(9), 6301–6310. https://doi.org/10.1016/j.apm.2013.01.010

Toloo, M. (2015). Alternative minimax model for finding the most efficient unit in data envelopment analysis. *Computers and Industrial Engineering*, *81*, 186–194. https://doi.org/10.1016/j.cie.2014.12.032

Toloo, M., & Salahi, M. (2018). A powerful discriminative approach for selecting the most efficient unit in DEA. *Computers and Industrial Engineering*, *115*, 269–277.

Wang, H. (2015). A generalized MCDA-DEA (multi-criterion decision analysis-data envelopment analysis) approach to construct slacks-based composite indicator. *Energy*, *80*, 114–122. https://doi.org/10.1016/j.energy.2014.11.051

Wang, Y. M., & Jiang, P. (2012). Alternative mixed integer linear programming models for identifying the most efficient decision making unit in data envelopment analysis. *Computers and Industrial Engineering*, *62*(2), 546–553.

https://doi.org/10.1016/j.cie.2011.11.003

Witte, K. De, & López-torres, L. (2017). Efficiency in education : a review of literature and a way forward. *Journal of the Operational Research Society*, *68*, 339–363. https://doi.org/10.1057/jors.2016.15

Yalçin, S., & Tavşancil, E. (2014). The Comparison of Turkish Students' PISA Achievement Levels by Year via Data Envelopment Analysis. *Kuram ve Uygulamada Egitim Bilimleri*, *14*(3), 947–968. https://doi.org/10.12738/estp.2014.3.1748

BIOGRAPHICAL SKETCH

Ece Uçar, born in Ankara in September 1996, is a research assistant at Galatasaray University Industrial Engineering Department. She graduated from Tevfik Fikret French High School in Ankara. She finished her Bachelor's Degree at Galatasaray University Industrial Engineering Department. She graduated second in success rank with a GPA of 3.76/4.00. She continued her education with a Master Program on Industrial Innovation at Grenoble Polytechnic Institute in France. During her master education, she was awarded with French Embassy Excellence Scholarship Program.

Now, she is about to finish her Master of Science Degree on Industrial Engineering at Galatasaray University Institute of Science. Her research areas are multi-criteria decision-making, efficiency analysis and education management. She plans to continue her career with Industrial Engineering PhD program at Galatasaray University Institute of Science.