

## Bootstrap Veri Zarflama Analizi ile TIMSS Verileri Kullanılarak Eğitim Sisteminde Teknik ve Yönetsel Etkinliklerin Karşılaştırılması

Serpil AYDIN<sup>\*1</sup>, Firdevs CEVHEROĞLU EREN<sup>2</sup>

<sup>1,2</sup> Ondokuz Mayıs Üniversitesi, Fen Edebiyat Fakültesi, İstatistik Bölümü, 55200, Samsun, Türkiye

(Alınış / Received: 26.05.2022, Kabul / Accepted: 27.02.2023, Online Yayınlanma / Published Online: 25.08.2023)

### Anahtar Kelimeler

Etkinlik,  
Bootstrap,  
Veri Zarflama,  
TIMSS,  
Eğitim

**Öz:** Araştırma, TIMSS 2015 kapsamında 4. sınıf matematik ve fen sınavına katılan 58 ülkenin başarıları ile ilişkili olduğu düşünülen girdilerini, etkin şekilde yönetip yönetmediklerini belirlemeyi amaçlamıştır. Bu amaçla, ülkelerin göreceli toplam etkinlik skorlarını tahmin etmek için parametrik olmayan yeniden örnekleme veri zarflama analizinden (bootstrap VZA) yararlanılmıştır. Bootstrap VZA yaklaşımının kullanılmasının nedeni, tahmin edilen sınırlı örnekleme varyasyonlarına göre etkinlik puanlarının duyarlılığını analiz etmenin kolay bir yolu olmasıdır. Bu çalışmada değerlendirilen veriler, öğrencilerin, öğretmenlerin ve okul müdürlerinin doldurduğu anketlerin IEA IDB Analyzer programıyla öğrencilerin matematik ve fen başarılarını temel alarak elde edilmiştir. Fen ve matematik başarısına etki eden eğitim girdilerini göreceli olarak etkin kullanan ülkeler sırasıyla Rusya, Amerika ve Çin- Taipei (Tayvan) olarak bulunmuş olup, bu bölgelerin neredeyse optimal düzeyde faaliyet gösterdiği ve istenen sabit çıktıyı üretmek için girdi tüketimlerini de neredeyse optimal kullandıkları gözlemlenmiştir. En düşük etkinliğe sahip olan ülkeler ise sırasıyla; Ermenistan, Karadağ, Birleşik Arap Emirlikleri (Dubai), Kuzey Makedonya, Kuveyt, Gürcistan ve Filipinler olarak belirlenmiştir. Türkiye ise etkin olmaya en yakın ülke olarak tespit edilmiştir.

## Comparison of Technical and Managerial Efficiencies in Education System Using Bootstrap Data Envelopment Analysis and TIMSS Dataset

### Keywords

Efficiency,  
Bootstrap,  
Data envelopment,  
TIMSS,  
Education

**Abstract:** The research aimed to determine whether 58 countries participating in the 4th grade mathematics and science exam within the scope of TIMSS 2015 effectively manage their inputs, which are thought to be related to their success. For this purpose, non-parametric resampling data envelopment analysis (bootstrap DEA) was used to estimate the relative total efficiency scores of countries. The reason for using the Bootstrap DEA approach is that the estimated bound is an easy way to analyze the sensitivity of event scores to sampling variations. The data used in this study were obtained based on the mathematics and science achievements of the students, with the IEA IDB Analyzer program of the questionnaires filled out by the students, teachers and school principals. Countries that use educational inputs that affect mathematics and science achievement relatively effectively are Russia, America and China-Taipei (Taiwan), respectively, and it has been observed that these regions operate at an almost optimal level and use their input consumption almost optimally to produce the desired stable output. The countries with the lowest efficiency are respectively; Armenia, Montenegro, United Arab Emirates (Dubai), North Macedonia, Kuwait, Georgia and the Philippines. Turkey, on the other hand, was determined as the country closest to being active.

### 1. Giriş

Eğitim sistemlerinin ve eğitim kurumlarının verimliliğinin değerlendirilmesi, günümüzde

hükümetlerin ve eğitim yetkililerinin tartışmalarının ön sıralarında yer almaktadır. Kamu bütçelerinin azalması, kamu harcamalarının sınırlı olmasından kaynaklı olarak kurumlar arasındaki rekabetin

artması, okulların faaliyetlerinin maliyetleri ve faaliyet sonuçlarına yönelik olarak talep edilen bilgideki şeffaflık gibi çeşitli faktörler verimliliği her zamankinden daha önemli hale getirmektedir. Dolayısıyla, eğitim kurumlarının verimliliği hakkında nicel bilgi sağlanması gerekmektedir. Bu bilgi, azalan kaynaklar karşısında bile daha iyi sonuçlarla ilişkili okul faktörlerinin daha iyi anlaşılmasına ve okul çıktılarında iyileştirmelerin desteklenmesine katkıda bulunabilir [1].

Öğrencilerin matematik ve fen alanlarında kazandıkları bilgi ve becerilerin başarı ölçütü olarak değerlendirilmesine yönelik bir araştırma yapan TIMSS (Trends in International Mathematics and Science Study-Uluslararası Matematik ve Fen Eğilimleri Araştırması), Uluslararası Eğitim Başarılarını Değerlendirme Kuruluşu'nun bir projesi olarak sunulmaktadır. TIMSS, dünyanın dört bir yanındaki öğrencilerin matematik ve fen bilgilerine ilişkin bir dizi uluslararası değerlendirmedir. TIMSS'in amacı, matematik ve fen öğretimi ve öğrenimini geliştirmeye yardımcı olmak için eğitim politikasıyla ilgili en iyi bilgileri sağlamaktır. Bu kapsamda TIMSS 2019, öğrenci, öğretmen ve okul anketlerini ve öğrencilerin ebeveynleri veya bakıcıları tarafından doldurulan Erken Öğrenme Anketini içermektedir. Anket maddelerinin çoğu, okulların önemli yönlerine (örneğin, okul güvenliği, öğrencilerin matematik ve fen öğrenmeye yönelik tutumları, erken başlamanın etkisi) ilişkin sağlam göstergeler sağlamak için madde yanıt teorisi ölçekleri olarak geliştirilmiş ve analiz edilmiştir. Anket sonuçları, öğrencilerin matematik ve fen öğrendiği ev, okul ve sınıf bağlamları hakkında zengin bilgiler sağlamaktadır. Değerlendirme kapsamındaki öğrenciler, ekonomik gelişme, coğrafi konum ve nüfus büyüklüğü açısından çeşitli ülkelerin sahip oldukları çeşitli eğitim sistemlerinden gelmektedir. Katılan eğitim sistemlerinin her birinde en az 4.000 ila 5.000 öğrenci değerlendirilir [2]. Bu değerlendirmeye ilk olarak 1995 yılında başlanmış ve her 4 yılda bir süreklilik göstermiştir. Bu nedenle, katılan eğitim sistemlerinden bazıları, 1995'ten 2019'a kadar olan değerlendirmeler arasında eğitim verilerine sahiptir [3]. TIMSS aracılığıyla 4. ve 8. sınıf öğrencileri, TIMSS Advanced aracılığıyla ise ortaokulun son yılındaki öğrencilerin ileri matematik ve fizik alanlarındaki başarı seviyeleri değerlendirilmektedir. Dört yıl arayla uygulanan TIMSS sınavları sonucunda ülkeler kendi gelişmelerini, uyguladıkları yeniliklerin amaçlarına ulaşip ulaşmadıklarını görebilmektedir. Yapılan sınavların ve değerlendirilen anketlerin verileri tüm araştırmacılara açık olup detaylı analizlerin yapılması için altyapı imkânı sağlamaktadır. Dolayısıyla bu şekilde yapılan sınavların ülkelerin kendi problemlerine hem özelde hem de uluslararası projeler aracılığıyla işbirliği yaparak çözüm aramalarına olanak sağladığı ifade edilebilir. Aynı zamanda evrensel düzeyde bilgi, beceri kazanacak insan kaynağını yetiştirmeye

yönelik farkındalık oluşturduğu; eğitim ve öğretim sürecinin güncellenmesine katkı sağladığı söylenebilir. Bu tür sınav sonuçları sayesinde ülkeler küresel ekonomide rakibi olan ülkelerin insan kaynakları kalitesini göreberek kendi politikalarına yön verebilmektedir. Bu da ülkeler arası politikalarda uzun vadede çok büyük avantajlar sağlayacaktır [4].

Uluslararası düzeyde eğitim verimliliği ile ilgili literatür, iki paralel akım üretmiştir. İlk kısım, analiz birimleri olarak ülkelere odaklanmaktadır. Diğer kısım ise okulları analiz birimi olarak değerlendirmektedir. Giménez vd. [5], 31 ülkedeki eğitim sisteminin verimliliğini analiz etmek için veri zarflama analizini kullanmıştır. Eğitim sisteminin verimliliğini, öğrencilerin ekonomik ve sosyal geçmişlerini dikkate alarak değerlendirmenin önemli olduğunu belirterek, eğitim sistemlerinin en verimli yönünün Komünist geçmişe sahip ülkelerde bulunabileceğini göstermişlerdir. Ayrıca, elde edilen sonuçlara göre, öğrencilerin performansını şu anda eğitim sistemlerine tahsis edildenden daha az kaynakla arttırabilecek bir dizi gelişmiş ülke olduğunu öne sürmektedirler. 2006–2009 döneminde 20 Avrupa ülkesinde eğitime yapılan harcama verimliliğini karşılaştıran Agasisti [6], çıktı olarak OECD-PISA test puanlarını, girdi olarak ise öğrenci başına harcama değişkenlerini kullanmıştır. Etkinlik sonuçları Bootstrap veri zarflama analizi kullanılarak elde edilmiştir. İkinci aşamada farklı sosyo – ekonomik durumların ve eğitim sisteminin yapısal özelliklerinin etkinlik üzerindeki etkisi araştırılmıştır. Çalışmanın sonucunda, öğretmenlerin maaşlarının ve internet kullanımının eğitim performansını olumlu yönde etkilediği ancak GSYİH'nın eğitim performansını olumsuz yönde etkilediği bulunmuştur. Yunanistan'da ilköğretim etkinliği düzeyini ölçmek ve okul ortamıyla (iç ve/veya dış) ilgili faktörlerin bu etkinlik düzeyini nasıl farklılaştırdığını incelemek isteyen Tsakiridou ve Stergiou [7] ilk aşamada VZA kullanarak etkinlik analizi yapmıştır ve ortalama verimliliği %76,26 bulmuştur. İkinci aşamada verimsizliğin sebeplerini açıklamak amacıyla sosyo ekonomik durumları ele almışlardır. Çalışma sonucunda öğrencilerin ailelerinin sosyo-ekonomik durumunun, okul alanının ve okul yeniliklerinin verimlilik ile pozitif ilişkili olduğunu; verimliliğe en büyük olumlu etkinin babanın eğitim düzeyinin yüksek olmasının, buna karşılık annenin eğitim düzeyinin düşük olmasının ise okulun verimliliğini olumsuz yönde etkilediğini bulmuşlardır. İslam konferansı teşkilatına üye 16 ülkenin ortaöğretim düzeyinde teknik verimlilik düzeyini araştırmayı amaçlayan Arshad [8], 40 ülkeyi kapsayan TIMSS 2011 verilerini ve teknik verimlilik düzeyini hesaplamak için veri zarflama analizi tekniğini kullanmıştır. Seçilen 16 İslam konferansı teşkilatı üyesinin neredeyse tamamı, teşkilat üyesi olmayan ülkelere kıyasla daha iyi TIMSS sonuçları elde etmek için eğitim kaynaklarını kullanmada teknik olarak yetersiz bulunmuştur. Çevresel

faktörleri kontrol ettikten sonra bile, islam konferansı teşkilat üyesi ülkelerde orta öğretim teknik olarak yetersiz kalmaktadır. Türkiye'deki sekizinci sınıf öğrencilerinin matematik başarısını etkileyen öğrenci özellikleri (akran zorbalığına ilişkin algı, öğrencilerin matematiğe olan güveni, öğrencilerin matematiği öğrenmeyi sevmeleri ve öğrencilerin matematiğe değer vermeleri) ile okul düzeyinde ele alınan öğretmen özellikleri arasındaki ilişkiyi araştıran Yavuz vd. [9], yapılan analiz sonucunda, okul düzeyinde, öğrencinin matematik başarısı ile öğretmenin akademik başarıya vurgusu değişkeni arasında pozitif ve anlamlı bir ilişki bulunmuştur. Öğrencilerin matematik başarıları ile öğretmenlerin çalışma koşulları ve öğretmenlerin öğretim değişkenlerini iyileştirmeye yönelik iş birliği arasında anlamlı bir ilişki bulunmamıştır. Öğrencilerin matematik sınavlarındaki performansları, öğrencilerin okulda zorbalığa maruz kalmamaları ve öğrencilerin matematik öğrenmeyi sevme değişkenleri ile öğrenci düzeyinde anlamlı bir ilişkiye sahip olduğu bulunmuştur. Agasisti ve Zoido [10], PISA 2012 verilerini kullanarak 30 ülkedeki yaklaşık 8500 okulun etkinliklerini veri zarflama analizi tekniğiyle hesaplamışlardır. Girdilerin sabit tutularak okulların başarı puanlarının %27 oranında artırılabilmesi çalışma sonucunda görülmektedir. Buna ek olarak, ilköğretim dördüncü sınıf öğrencilerinin 2011 yılına ait TIMSS verilerini kullanarak Fas'taki ilköğretim okullarının verimliliğini analiz etmek için Liouaeddine vd. [11] teknik olarak Veri Zarflama Analizi kullanmıştır. Çalışma sonucunda, öğrencilerin sosyo-ekonomik ortamının Fas'taki ilkokulların verimliliği üzerinde önemli bir etkisi olduğu ortaya çıkmıştır. Koyuncu ve Ilgaz [4], TIMSS 2015 kapsamında 8. sınıf matematik çalışmasına katılan 41 ülkenin matematik başarısıyla ilişkili olduğu düşünülen girdilerini etkin şekilde yönetip yönetmediklerini belirlemeyi amaçlamış ve verilerin analizinde veri zarflama analizinden faydalanılmıştır. Çalışma sonucunda düşük etkinliğe sahip ülkelerin kaynaklarını verimsiz kullandıklarını ve harcamalarının karşılığını tam olarak alamadıklarını saptamışlardır. Mazurek ve Mielcová [12], 15 ve 16 yaşındaki öğrencilerin sosyo ekonomik göstergeleri ile PISA sonuçları arasındaki ilişkiyi incelemişlerdir. 34 OECD ülkesi ve 37 OECD üyesi olmayan 71 ülkeyi kapsayan araştırmanın temel sonucu, kişi başına düşen GSYİH ve kişi başına ilköğretime yapılan devlet harcamaları açısından bir eşik bulunmasıdır. Eşiğin üzerinde, kişi başına düşen GSYİH veya harcamalar, daha yüksek PISA puanlarına dönüşmez. Ancak, bu eşiğin altında, bunun tersi doğrudur. Bu nedenle, daha yoksul ve çoğunlukla OECD üyesi olmayan ülkeler, ilköğretime yapılan harcamaları artırarak PISA testlerinde daha iyi öğrenci performansları elde edebilirken, zenginlerin ve ağırlıklı olarak OECD ülkelerinin öğrenci performansları için harcamalar istatistiksel olarak önemli bir faktör değildir. Haddad vd. [13], eğitim verimliliğini değerlendirmek için kullanılan teknikler

üzerinde durarak, mevcut literatürü sentezleyerek eğitim verimliliği hakkında bilgi sağlamaktadır. "Web of Science" "Scopus" ve "Google Scholar"ın ilgili sistematik veri tabanlarında "Education Efficiency", "Okul Performansı" ve "Efficiency Measurements" terimleri kullanılarak mevcut literatürü kapsamlı bir şekilde taramışlardır. Sonuçlara bakıldığında, eğitimin kilit paydaşlarına önemli bilgiler ve avantajlar sağlaması açısından VZA'nın eğitimin en önemli ölçümlerini yönlendirmek için literatürde kullanılan birincil metodolojilerden biri olduğunu gösterdiler.

Parametrik olmayan bir çerçeve içerisinde üretim etkinlik skorlarını tahmin etme fikri, ilk olarak Farrell [14]'in çalışmasına dayanmaktadır. Üretim birimlerinin etkinliği, tipik olarak, optimal üretim planlarının geometrik yeri olarak tanımlanan bir üretim sınırına göre ölçülmektedir. Bu sınır, bir dizi n adet gözlemlenen üretim biriminden (parametrik olmayan) tahmin edilebilmektedir. VZA modelleri Charnes vd. [15]'nin yapmış olduğu çalışma ile ortaya çıktıktan sonra ilk olarak Amerika'daki eğitim politikalarının analizinde uygulanmıştır. VZA yönteminin ana avantajlarından biri, verimli hedefler kümesi belirleme ve optimizasyon analiziyle benzerlerini belirleme olasılığıdır. Bu sıralamaların oluşturulması için daha uygun olan sabit ağırlıklara dayalı geleneksel yöntemlere göre büyük bir avantajdır. Geleneksel yöntemler daha çok performans yönetimine odaklanırken, VZA daha çok iyileştirme arayışını içeren performans yönetimine odaklanmıştır [16]. Bu yöntemde sınır, ölçüğe göre değişken getirilere izin verirken, girdi/çıktı uzayındaki gözlemler kümesinin dışbükey gövdesinin sınırı olarak alınır. Ayrıca bu yaklaşım, ulaşılabilir üretim planları setinin dışbükeylik varsayımlarına dayanmaktadır [17;18]. Efron [19] tarafından tanımlanan yeniden örnekleme (bootstrap), ölçülen etkinlik skorlarının örnekleme varyasyonuna duyarlılığını analiz etmek için faydalı bir araç olarak görülmüştür. Bootstrap, genellikle yeniden örnekleme yoluyla veri üretme sürecini tekrar tekrar simüle etme ve orijinal tahmin ediciyi simüle edilmiş her bir örneğe uygulama fikrine dayanır. Böylece sonuç tahminleri orijinal tahmin edicinin örnekleme dağılımını taklit eder.

Bu çalışma, literatürün ikinci kısmına yani okulların analiz birimi olarak değerlendirildiği durumda eğitim çıktılarını iyileştirmek amacıyla okulların nasıl organize edilebileceğini veya kaynak sağlanabileceğini analiz etmektedir. Dolayısıyla, TIMSS 2019 kapsamında 4. sınıf matematik ve fen çalışmasına katılan 58 ülkenin başarısıyla ilişkili olduğu düşünülen girdilerini etkin şekilde yönetip yönetmediklerinin belirlenmesi amaçlanarak, nicel araştırma kapsamında yer alan betimsel tarama modeli kullanılmıştır. Verilerin analizinde ise bootstrap VZA tekniğinden faydalanılmıştır. Bu çalışmayı özgün kılan nedenlerden birisi de ülkelere

ve kurumlara görece olarak bilgi sunan veri zarflama analizinin yanlılığını da ortadan kaldıran bootstrap veri zarflama analizi kullanılmış olmasıdır. Böylece ülkeler uluslararası çalışmalar ile elde edilen verileri kullanarak eğitim ile ilgili emeklerinin, projelerinin ve girdilerinin şu anki durumlarını, geleceğe yansımalarını, hedeflerinin gerçekleşme düzeylerini ve diğer ülkeler ile kendilerini karşılaştırma yaparak Dünya'daki konumlarını görebilmektedirler.

## 2. Materyal ve Metot

### 2.1. Veri zarflama analizi

Veri zarflama analizi, belirli girdileri ve belirli çıktıları kullanarak üretim sınırlarının tahmini için çalışmalarda sıklıkla başvurulan, 1978 yılında Charnes, Cooper ve Rhodes tarafından ölçeğe göre sabit getirili modeli ve ardından 1984 yılında Banker, Charnes ve Cooper tarafından ölçeğe göre değişken getirili modeli geliştirilen karar verme birimleri için görece etkinlik kıyaslaması yapan parametrik olmayan bir yöntemdir [20; 21]. Karar verme birimlerinin etkinliklerinin ölçülmesinde kullanılır. Her bir karar verme birimi için iki durum söz konusudur: etkin veya etkin değil. Karar verme biriminin etkinlik seviyesinin belirlenmesi ağırlıklı çıktıların toplamının, ağırlıklı girdilerinin toplamına oranlanması ile gerçekleştirilir. Karar verme biriminin işletilmesinde etkinsiz süreç varsa ya da çalışma şartları içerisinde dezavantajlı bir duruma

sahip olursa karar verme birimi etkinsiz olur [22]. Bunu belirleyebilmek için de CRS etkinlik skorunun VRS etkinlik skoruna oranı ile elde edilen ölçek etkinliğine başvurulur. VZA modelleri temel olarak girdi ve çıktı yönlü olmak üzere iki ana gruba ayrılabilirler. Girdi yönlü olanlar, herhangi bir çıktı seviyesine ulaşmak için etkin olmayan karar birimlerinin girdilerini ne derece azaltmaları gerektiğini araştırırlar. Yani sabit çıktı düzeyinde girdi kullanımını oransal olarak azaltarak teknik etkinsizliği ölçmeye çalışırlar. Benzer şekilde, çıktı yönlü modeller ise herhangi bir girdi düzeyinde kalabilmek için etkin olmayan karar birimlerinin etkin duruma getirilebilmesi amacıyla çıktıları ne kadar artırabilecekleri üzerinde dururlar. Dolayısıyla sabit girdi düzeyinde çıktı kullanımını oransal olarak artırarak teknik etkinsizliği ölçerler [23]. Girdi (çıkıtı) odaklı primal (dual) modeller sırasıyla Şekil 1 ve Şekil 2'deki gibi gösterilmektedir.

Burada  $x_{ij} \subseteq X$  ve  $X, n \times m$  boyutunda tüm karar verme birimleri için girdileri veren matris ve  $y_{rj} \subseteq Y$  ve  $Y, n \times s$  boyutunda tüm karar verme birimleri için çıktıları veren matristir.  $X_{io}$  ve  $Y_{ro}$  o. karar verme biriminin girdi ve çıktıları veren sırasıyla  $m \times 1$  ve  $s \times 1$  boyutlu vektörler;  $u_r$  ve  $v_i$  sırasıyla çıktıların ve girdilerin ağırlık vektörleridir. Yine bir skaler değer olan  $\theta_o$  girdi odaklı dual modelin ve  $\phi_o$  çıktı odaklı dual modelin o. karar verme biriminin etkinlik skorudur.

	Primal	Dual
<b>Girdi odaklı</b>	$\text{mak } z_o = \sum_{r=1}^s u_r y_{ro} - (u_o)^*$ <p><b>Kısıtlar:</b></p> $\sum_{i=1}^m v_i x_{io} = 1$ $-\sum_{i=1}^m v_i x_{ij} + \sum_{r=1}^s u_r y_{rj} - (u_o)^* \leq 0$ $v_i \geq 0 \quad u_r \geq 0 \quad u_o \text{ sınırsız}$	$\text{min } \theta_o$ <p><b>Kısıtlar:</b></p> $\sum_{j=1}^n \lambda_j x_{ij} - \theta_o x_{io} \leq 0 \quad i = 1, \dots, m$ $y_{ro} - \sum_{r=1}^s \lambda_j y_{ir} \leq 0 \quad r = 1, \dots, s$ $\left( \sum_{j=1}^n \lambda_j = 1 \right)^*$

Şekil 1. Girdi Odaklı Primal/Dual VZA Modelleri [15,17]

	Primal	Dual
<b>Çıkıtı odaklı</b>	$\text{min } z_o = \sum_{i=1}^m v_i x_{io} - (v_o)^*$ <p><b>Kısıtlar:</b></p> $\sum_{r=1}^s u_r y_{rj} = 1$ $\sum_{i=1}^m v_i x_{ij} - \sum_{r=1}^s u_r y_{rj} - (v_o)^* \leq 0$ $v_i \geq 0 \quad u_r \geq 0 \quad v_o \text{ sınırsız}$	$\text{mak } \phi_o$ <p><b>Kısıtlar:</b></p> $\sum_{j=1}^n \lambda_j x_{ij} - x_{io} \leq 0 \quad i = 1, \dots, m$ $\phi_o y_{ro} - \sum_{r=1}^s \lambda_j y_{ir} \leq 0 \quad r = 1, \dots, s$ $\left( \sum_{j=1}^n \lambda_j = 1 \right)^*$

(\*) BCC modelde kısıt olarak eklenir.

Şekil 2. Çıkıtı Odaklı Primal/Dual VZA Modelleri [15,17]

## 2.2. Bootstrap veri zarflama analizi

VZA, KVB'lerinin üretim sürecinde kullandığı girdi ve çıktı değişkenlerine optimal ağırlıklar verebilmek için doğrusal programlama kullanır. VZA'nın en büyük avantajı parametrik olmayan bir yaklaşım olduğu için belirli bir üretim fonksiyonu gerektirmemesidir. Öte yandan en büyük dezavantajı da deterministik yapısı sebebiyle istatistiksel çıkarsama yapmanın mümkün olmamasıdır. Bu dezavantajı azaltmak için Simar ve Wilson [24], rastgeleliğe karşı VZA skorlarının duyarlılığını ortaya çıkaran Bootstrap VZA yöntemini önerdiler. Tahmin edicilerin örnekleme dağılım özelliğinden dolayı geleneksel nokta tahminleri, VZA'yı tutarlı bir etkinlik tahmincisi olarak görmek için yetersizdir [25]. Bu sınırlamanın üstesinden gelmek için Efron [19] yeniden örnekleme (bootstrap) metodolojisi, tahmini sınırın örnekleme varyasyonlarına göre ölçülen etkinlik puanlarının duyarlılığını değerlendirmek için uygun bir araç olarak karşımıza çıkmaktadır. Simar ve Wilson [24], geleneksel VZA prosedüründen daha teknik etkinlik tahminlerinin güvenilirliğini artırmak için Efron [19]'un yaklaşımlarına dayanan parametrik olmayan sınır modelleri için genel bir metodoloji ortaya koymuştur. Ayrıca, basitleştirilmiş bir şekilde, bootstrap'ın, tipik olarak orijinal tahmin edicinin örnekleme dağılımını çoğaltmak için yeniden örnekleme ve orijinal tahmin ediciyi her benzetilmiş örneğe ekleyerek, veri üretme sürecini tekrar tekrar simüle etme ilkesine dayandığını belirtmişlerdir. Bu nedenle, nokta tahminlerinin örnekleme dağılımında stokastik hatalar varsa, analitik olarak türetilmeyen güven aralıklarını oluşturmak için bootstrap VZA prosedürü gerçekleştirilebilir [26].

Bu çalışmada Simar ve Wilson tarafından geliştirilen bootstrap VZA yöntemi uygulanmaktadır [24;27;28]. Bootstrapping, değişkenlerin dağılımı veya veri oluşturma süreci hakkında önceden varsayımlar veya bilgi olmadan etkinlik skorları için çıkarımsal istatistikler geliştirerek VZA'nın parametrik olmayan doğasına ideal bir çözüm sunmaktadır. Bootstrap VZA için, örnekleme dağılımları Monte Carlo simülasyonu kullanılarak yaklaştırılır.

Gerçek veri oluşturma sürecinin bilinmeyen P olduğunu varsayalım. Simar ve Wilson [24]'a göre, ampirik örnekleme dağılımına dayalı Monte Carlo simülasyonu kullanılarak,  $\bar{P}$  olarak adlandırılan tutarlı bir P tahmini, Eşitlik (1)'deki ilişki ile  $\bar{\theta}^*$  etkinlik skorlarına karşılık gelen bir vektörü verecek şekilde simüle edilebilir:

$$(\bar{\theta}^* - \hat{\theta})|\bar{P} \approx (\bar{\theta}^* - \hat{\theta})|P \quad (1)$$

Burada,  $\hat{\theta}$ , VZA'dan elde edilen orijinal etkinlik skoru olup,  $\theta^*$  etkinlik skorlarının gerçek ve bilinmeyen vektörüdür. Eşitlik (1) VZA ( $\hat{\theta}$ ) ile etkinlik puanının yanlılığını tahmin etmemizi sağlar. Simar ve Wilson [24]'e göre,  $\theta^*$  tahmini Eşitlik (2) ile hesaplanabilir.

$$\bar{\theta}^* = 2\hat{\theta} - \bar{\theta}^* \quad (2)$$

Burada,

$$\bar{\theta}^* = \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b \quad (3)$$

$$s\hat{e} = \sqrt{\frac{1}{B-1} \sum_{b=1}^B (\hat{\theta} - \bar{\theta}^*)^2} \quad (4)$$

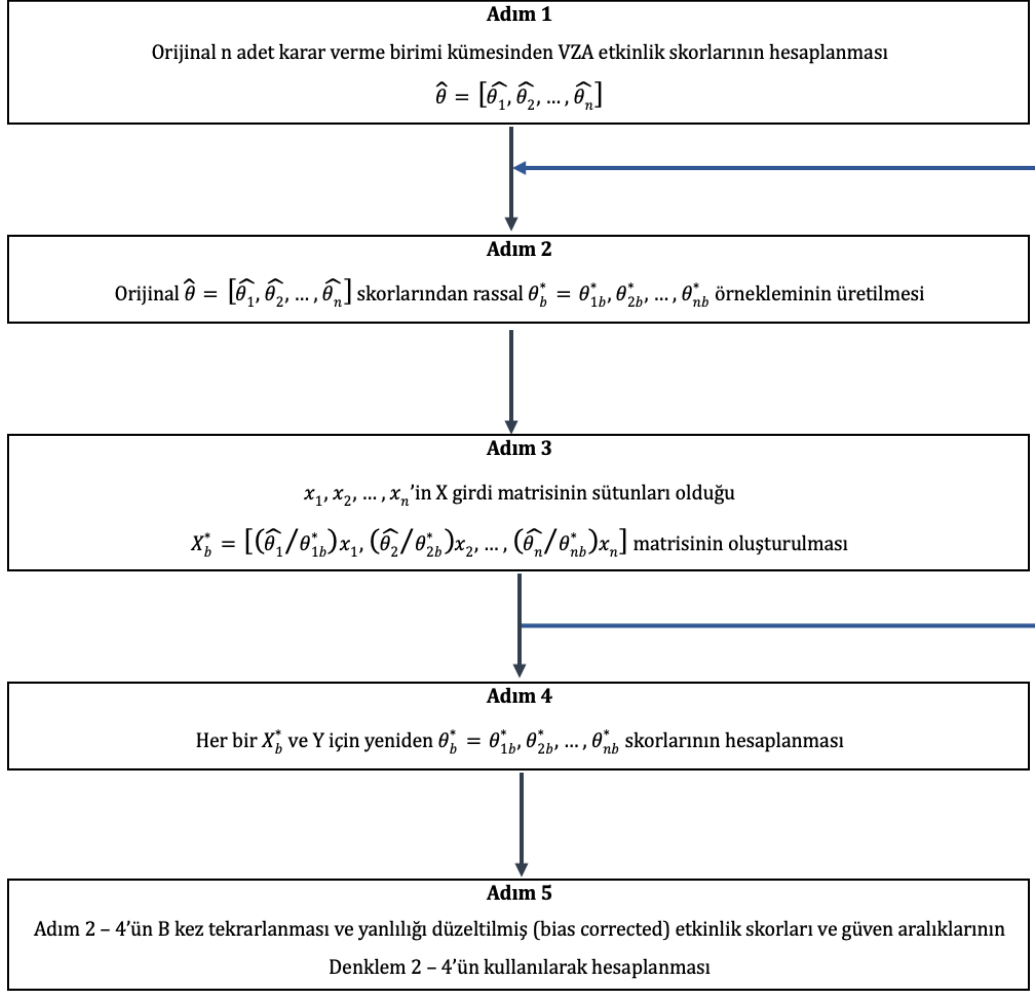
Yeniden örnekleme örneklerinin sayısına göre  $b = 1, 2, \dots, B$  'dir.  $\hat{\theta}_b$  'nin ampirik dağılımı çarpıksa, medyan değer kullanılabilir. Şekil 3, bootstrapped VZA skorlarında kullanılacak genel adımları göstermektedir. Adım 2, bootstrapping'in anahtarıdır. Simar ve Wilson [24], 1'lik üst sınıra yakın etkinlik puanlarının tahminini iyileştirmek için düzgün yeniden örnekleme (smooth bootstrap) yönteminin kullanılması gerektiğini önermektedir. 1000 ila 2000 kez yeniden örnekleme yapmak yaygın bir uygulamadır. Bu çalışmada, veri seti 2000 kez yeniden örneklendirilmiştir.

## 3. Bulgular

Literatürdeki araştırmalardan Koyuncu ve Ilgaz [4]'ün Tablo 1.'deki etkinlik göstergeleri girdi ve çıktı değişkenleri olarak referans alınmıştır. Bootstrap VZA metodolojisinde girdi odaklı CCR modelinin uygulanması tercih edilmiştir. Problemin çözümünde MaxDEA Ultra-8 yazılımından faydalanılmıştır. 58 karar verme birimi, 8 girdi ve 2 çıktı ile oluşturulan modelin çözümünden elde edilen etkinlik skorları Tablo 2.'de gösterilmektedir.

**Tablo 1.** Girdi ve çıktı değişkenleri

Girdiler
Öğretmenlik kıdemi
Akademik başarıda okulun etkisi
Okulun güvenli ve düzenli olması
Öğrenci ihtiyaçları ile sınırlandırılmış öğretim
Evdeki eğitim kaynakları
Matematik öğretim kaynağı ile öğretim sıkıntısı
Akademik başarıda okulun etkisi
Okul disiplin problemleri
Fen bilimleri öğretim kaynağı ile öğretim sıkıntısı
Çıktılar
Ortalama matematik başarısı
Ortalama bilim başarısı



Şekil 3. Bootstrap VZA Adımları [24]

Tablo 2. Etkinlik Skorlarının Duyarlılık Analizi

Ülkeler	Etkinlik Skoru	Ortalama	Medyan	Alt Sınır	Üst Sınır
Almanya	0,9636	0,9426	0,9403	0,9240	0,9719
Amerika	1,0000	0,9748	0,9754	0,9524	1,0009
Arnavutluk	0,9106	0,8859	0,8774	0,8632	0,9399
Avustralya	0,9466	0,9290	0,9287	0,9141	0,9475
Avusturya	0,9299	0,9066	0,9001	0,8857	0,9509
Azerbaycan	0,8862	0,8605	0,8586	0,8375	0,8975
Bahreyn	0,9200	0,8911	0,8806	0,8647	0,9671
Belçika	0,9506	0,9276	0,9262	0,9074	0,9580
Birleşik Arap Emirlikleri	0,9116	0,8883	0,8858	0,8675	0,9215
Birleşik Arap Emirlikleri (Abu Dhabi)	0,8788	0,8559	0,8531	0,8355	0,8898
Birleşik Arap Emirlikleri (Dubai)	0,8505	0,8239	0,8167	0,7997	0,8782
Bulgaristan	0,8651	0,8468	0,8455	0,8312	0,8667
Çek Cumhuriyeti	0,9964	0,9719	0,9692	0,9501	1,0032
Çin-Taipei	0,9970	0,9739	0,9715	0,9535	1,0023
Danimarka	0,9910	0,9693	0,9694	0,9505	0,9912
Ermenistan	0,8430	0,8301	0,8290	0,8194	0,8462

**Tablo 2.** Etkinlik Skorlarının Duyarlılık Analizi (Devam)

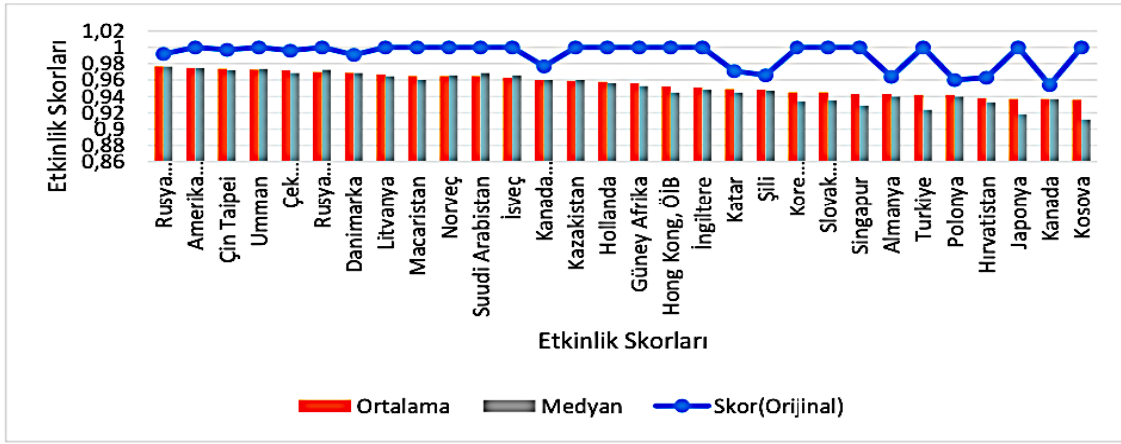
Ülkeler	Etkinlik Skoru	Ortalama	Medyan	Alt Sınır	Üst Sınır
Fas	0,9044	0,8820	0,8795	0,8623	0,9157
Filipinler	0,5933	0,5806	0,5801	0,5695	0,5963
Fransa	0,9448	0,9192	0,9192	0,8964	0,9476
Güney Afrika	1,0000	0,9564	0,9534	0,9155	1,0102
Gürcistan	0,8050	0,7870	0,7867	0,7712	0,8063
Hırvatistan	0,9630	0,9376	0,9325	0,9149	0,9886
Hollanda	1,0000	0,9575	0,9561	0,9177	1,0237
Hong Kong, SAR	1,0000	0,9520	0,9451	0,9071	1,0223
İngiltere	1,0000	0,9514	0,9479	0,9057	1,0237
İspanya	0,8535	0,8349	0,8340	0,8186	0,8578
İspanya, Madrid	0,8599	0,8397	0,8395	0,8220	0,8614
İsveç	1,0000	0,9634	0,9661	0,9298	1,0007
Japonya	1,0000	0,9371	0,9179	0,8770	1,0498
Kanada	0,9543	0,9367	0,9366	0,9214	0,9541
Kanada (Ontario)	0,9774	0,9600	0,9598	0,9454	0,9771
Kanada (Quebec)	0,9581	0,9311	0,9286	0,9072	0,9707
Katar	0,9705	0,9487	0,9454	0,9297	0,9846
Kazakistan	1,0000	0,9589	0,9599	0,9207	1,0115
Kıbrıs	0,9454	0,9260	0,9258	0,9093	0,9453
Kore	1,0000	0,9453	0,9338	0,8934	1,0296
Kosova	1,0000	0,9361	0,9116	0,8750	1,0918
Kuveyt	0,8134	0,7906	0,7850	0,7700	0,8325
Kuzey Makedonya	0,8376	0,8149	0,8126	0,7945	0,8469
Letonya	0,9405	0,9187	0,9176	0,8994	0,9454
Litvanya	1,0000	0,9675	0,9652	0,9375	1,0096
Macaristan	1,0000	0,9645	0,9599	0,9319	1,0265
Malta	0,9495	0,9283	0,9258	0,9099	0,9588
Montenegro	0,8508	0,8286	0,8271	0,8089	0,8563
Norveç	1,0000	0,9648	0,9662	0,9322	1,0096
Pakistan	0,9330	0,9151	0,9142	0,8993	0,9370
Polonya	0,9599	0,9419	0,9402	0,9262	0,9706
Portekiz	0,9474	0,9274	0,9265	0,9101	0,9510
Rusya	1,0000	0,9697	0,9729	0,9421	0,9969
Rusya, Moskova	0,9916	0,9767	0,9765	0,9646	0,9897
Sırbistan	0,9282	0,9084	0,9057	0,8910	0,9383
Singapur	1,0000	0,9430	0,9287	0,8888	1,0305
Slovakya	1,0000	0,9454	0,9348	0,8935	1,0355
Suudi Arabistan	1,0000	0,9651	0,9695	0,9331	0,9997
Şili	0,9655	0,9477	0,9471	0,9326	0,9682
Türkiye	1,0000	0,9418	0,9235	0,8863	1,0532
Umman	1,0000	0,9730	0,9736	0,9486	1,0020
Yeni Zelanda	0,9322	0,9111	0,9113	0,8928	0,9317

Bootstrap VZA sonuçları etkinlik ölçümlerinin örnekleme varyasyonuna karşı duyarlılığını ortaya koymaktadır. Tablo 2.'deki sonuçlara göre, orijinal VZA etkinlik skorlarına dayalı ülkelerarası göreceli performans karşılaştırmasında dikkatli olunması gerektiği görülmektedir. Örneğin Çin Taipei'nin orijinal VZA etkinlik skoru 0,997'dir, İngiltere ise 1,000 etkinlik skoru ile görünüşte etkin gözükmemektedir. Bununla birlikte, iki ülkenin etkinliği için güven aralıklarının büyük ölçüde örtüştüğü görülmektedir. Bu nedenle, iki ülkenin toplam etkinlikleri açısından önemli ölçüde fark olduğunu söyleyemeyiz.

Analiz edilen dönemde ortalama etkinlik skorları sıralamasına göre göreceli olarak en etkin ülke Rusya (0,977), onu Amerika (0,975) ve Çin-Taipei

(Tayvan) (0,974) takip etmekte olup, bu bölgelerin neredeyse optimal düzeyde faaliyet gösterdiği ve istenen sabit çıktıyı üretmek için girdi tüketimleri de neredeyse optimal olduğu anlamına gelmektedir. Ayrıca, bu sonuçlar, bu üç bölgenin, kaynak tasarrufu yaklaşımını izleyerek arzu edilen sabit çıktı düzeyine ulaşmak için mevcut kaynaklarını etkin bir şekilde yönettiklerinin varsayıldığını göstermektedir.

Bununla birlikte geleneksel VZA skorlarına göre sadece Amerika etkin (1,00) olup diğer iki ülke etkisiz çıkmıştır. Benzer şekilde söz konusu geleneksel VZA skoruna göre Türkiye Almanya'ya göre etkin çıkmışken Bootstrap VZA yöntemine göre Almanya'nın Türkiye'ye göre kaynaklarını daha etkin kullandığını göstermektedir.



Şekil 4. Ortalama Etkinlik Skorlarına göre sıralanmış ülkelerin VZA skor değerleri

#### 4. Tartışma ve Sonuç

Bu çalışma, parametrik olmayan Bootstrap VZA yöntemini kullanarak sınav başarısına yönelik kaynak verimliliği yaklaşımını değerlendirmiştir. Bu durum değerlendirmesi, politika yapıcılara ve planlayıcılara, bu politika yaklaşımının güçlü ve zayıf yönlerini ve bu maliyetli politikayı sürdürmenin önemini belirlemede yardımcı olabileceğini sağlayacak bir tavsiye niteliğindedir. Üstelik, Bootstrap VZA yaklaşımı, geleneksel VZA prosedürü kullanılarak elde edilenlerden istatistiksel olarak daha doğru verimlilik puanları ile sonuçlanan, geleneksel verimlilik puanlarını düzeltmemize ve güven aralıkları elde etmemize izin vermiştir. Türkiye dördüncü sınıflarda 58 ülke arasından fen sınavı bakımından 36'ncı, matematik sınavı bakımından 22'inci ve her iki sınavın ortalamaları bakımından 26. sırada yer almaktadır. Türkiye'nin sınav başarı sonuçları değerlendirildiğinde etkin olmaya en yakın ülke olmasına rağmen alt sıralarda yer almaktadır. Değişkenlerden elde edilen bulguların gerçeği yansıttığı düşünüldüğünde sonuç olarak, özellikle tam etkin olmayan ülkelerin çoğunun eğitime ihtiyaçlarından fazla kaynak ayırdıkları veya

girdilerin başarıya dönüşümü konusunda etkin olamadıkları söylenebilir. Ermenistan, Karadağ, Birleşik Arap Emirlikleri (Dubai), Kuzey Makedonya, Kuveyt, Gürcistan, Filipinler gibi az gelişmiş veya gelişmekte olan ülkeler eğitime yönelik yaptıkları harcamaların, başarıya odaklı dönüşümünü sağlama açısından daha fazla çaba göstermeli veya eğitime ayırdıkları bütçenin ihtiyaçtan fazla olan kısmında girdileri azaltma yoluna gitmelidir. Çalışma, bölgelerin etkinlik sıralamasında tek aşamalı Bootstrap VZA yöntemini uygulamaktadır. Bununla birlikte, verimlilik tahminindeki tutarlılığı artırmak için, gelecekteki araştırmalar Simar ve Wilson [29] tarafından geliştirilen iki aşamalı Bootstrap VZA yöntemini uygulamayı düşünebilir.

#### Etik Beyanı

*Bu çalışmada, "Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesi" kapsamında uyulması gerekli tüm kurallara uyulduğunu, bahsi geçen yönergenin "Bilimsel Araştırma ve Yayın Etiğine Aykırı Eylemler" başlığı altında belirtilen eylemlerden hiçbirinin gerçekleştirilmediğini taahhüt ederiz.*



## Kaynakça

- [1] Agasisti, T., & Zoido, P. 2015. The efficiency of secondary schools in an international perspective: preliminary results from PISA 2012.
- [2] IEA. 2021. International Association for the Evaluation of Educational Achievement. Retrieved from <https://timss2019.org/reports/about/>
- [3] PRILS, T. 2021. Trends in International Mathematics and Science Study. Retrieved from <https://timssandpirls.bc.edu/>
- [4] Koyuncu, B., ve Ilgaz, G. 2019. Matematik Öğretimi Sürecinde Ülkelerin Eğitim Girdilerini Ne Kadar Etkin Kullandıklarının TIMSS 2015 Verilerine Göre İncelenmesi. *İlköğretim Online*, 18(4).
- [5] Giménez, V., Prior, D., & Thieme, C. 2007. Technical efficiency, managerial efficiency and objective-setting in the educational system: an international comparison. *Journal of the Operational Research Society*, 58(8), 996-1007.
- [6] Agasisti, T. 2014. The efficiency of public spending on education: An empirical comparison of EU countries. *European Journal of Education*, 49(4), 543-557.
- [7] Tsakiridou, H., & Stergiou, K. 2014. Explaining the efficiency differences in primary school education using data envelopment analysis. *Journal of Education, Psychology and Social Sciences*, 2(2), 89-96.
- [8] Arshad, M. N. M. 2014. Efficiency of Secondary Education of a Selected OIC Countries. *Global Education Review*, 1(4).
- [9] Yavuz, H., Demirtasli, R., Yalcin, S., & Dibek, M. 2017. The effects of student and teacher level variables on TIMSS 2007 and 2011 mathematics achievement of Turkish students. *EGITIM VE BILIM-EDUCATION AND SCIENCE*, 42(189).
- [10] Agasisti, T., & Zoido, P. 2018. Comparing the efficiency of schools through international benchmarking: Results from an empirical analysis of OECD PISA 2012 data. *Educational Researcher*, 47(6), 352-362.
- [11] Liouaeddine, M., Elatrachi, M., & Karam, E. M. 2018. The analysis of the efficiency of primary schools in Morocco: modelling using TIMSS database (2011). *The journal of North African studies*, 23(4), 624-647.
- [12] Mazurek, J., & Mielcová, E. 2019. On the relationship between selected socio-economic indicators and student performances in the PISA 2015 study.
- [13] Haddad, M. Z., Heong, Y. M., Razzaq, A. R. B. A., & Kiong, T. T. 2021. Exploring the Innovative Methods for Evaluating Educational Efficiency. In 2021 International Conference on Decision Aid Sciences and Application (DASA)(pp. 1082-1086). IEEE.
- [14] Farrell, M. J. 1957. The measurement of productive efficiency. *Journal of the Royal Statistical Society: Series A (General)*, 120(3), 253-281.
- [15] Charnes, A., Cooper, W. W., & Rhodes, E. 1981. Evaluating program and managerial efficiency: an application of data envelopment analysis to program follow through. *Management science*, 27(6), 668-697.
- [16] Stumbriene, D., Camanho, A. S., & Jakaitiene, A. 2020. The performance of education systems in the light of Europe 2020 strategy. *Annals of Operations Research*, 288(2), 577-608.
- [17] Charnes, A., Cooper, W. W., ve Rhodes, E. 1978. Measuring the efficiency of decision making units. *European journal of operational research*, 2(6), 429- 444.
- [18] Fare, R., Grosskopf, S., ve Kokkelenberg, E. C. 1989. Measuring plant capacity, utilization and technical change: a nonparametric approach. *International economic review*, 655-666.
- [19] Efron, B. 1992. Bootstrap methods: another look at the jackknife. In *Breakthroughs in statistics* (pp. 569-593): Springer.
- [20] Førsund, F. R., ve Sarafoglou, N. 2002. On the origins of data envelopment analysis. *Journal of Productivity Analysis*, 17(1), 23-40.
- [21] Cooper, W. W., Seiford, L. M., ve Zhu, J. 2011. Data envelopment analysis: History, models, and interpretations. In *Handbook on data envelopment analysis* (pp. 1- 39): Springer.
- [22] Kutlar, A., ve Babacan, A. 2008. Türkiye'deki kamu üniversitelerinde CCR etkinliği- ölçek etkinliği analizi: DEA tekniği uygulaması. *Kocaeli Üniversitesi Sosyal Bilimler Dergisi*(15), 148-172.
- [23] Coelli, T., ve Perelman, S. 2000. Technical efficiency of European railways: a distance function approach. *Applied economics*, 32(15), 1967-1976.
- [24] Simar, L., ve Wilson, P. W. 1998. Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management science*, 44(1), 49- 61.
- [25] Simar, L. 1992. Estimating efficiencies from frontier models with panel data: a comparison of parametric, non-parametric and semi-parametric methods with bootstrapping. In *International Applications of Productivity and Efficiency Analysis* (pp. 167-199): Springer.

- [26] Toma, P., Miglietta, P. P., Zurlini, G., Valente, D., ve Petrosillo, I. 2017. A non-parametric bootstrap-data envelopment analysis approach for environmental policy planning and management of agricultural efficiency in EU countries. *Ecological indicators*, 83, 132-143.
- [27] Simar, L., ve Wilson, P. W. 2000. A general methodology for bootstrapping in non-parametric frontier models. *Journal of applied statistics*, 27(6), 779-802.
- [28] Simar, L., ve Wilson, P. W. 2002. Non-parametric tests of returns to scale. *European journal of operational research*, 139(1), 115-132.
- [29] Simar, L., ve Wilson, P. W. 2007. Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of econometrics*, 136(1), 31-64.