European Journal of Operational Research 305 (2023) 356–372

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Decision Support

The assessment of performance trends and convergence in education and training systems of European countries

Ana S. Camanho[a,*], Dovile Stumbriene[b], Flávia Barbosa[a], Audrone Jakaitiene[b]

[a] Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, Porto, 4200-465, Portugal
[b] Institute of Data Science and Digital Technologies, Vilnius University, Akademijos st. 4, Vilnius, LT-08412, Lithuania

ARTICLE INFO

Article history:
Received 29 November 2021
Accepted 30 May 2022
Available online 3 June 2022

Keywords:
Data envelopment analysis
European education systems
Composite indicators
Global malmquist index
Sigma and beta convergence

ABSTRACT

The Strategic Framework for European Cooperation in Education and Training (ET 2020) aimed to promote the exchange of best practices among the Member States. This paper assesses the performance evolution of European countries in terms of the common objectives for the education sector. The framework used to evaluate European education systems is based on constructing a composite indicator adopting a "benefit-of-the-doubt" approach. The evaluation of performance change over time is done using a Global Malmquist Index. Sigma and beta convergence of EU countries are also explored using non-parametric frontier techniques. The results are analysed for the period 2009–2018 and discussed in light of the goals envisaged and the national policies adopted. The results revealed a trend of improvement in the performance of education systems in most European countries in the period analysed. Although most European countries moved closer to the European best practice frontier over time, as confirmed by the values of sigma-convergence, a few countries are still lagging considerably below their peers, as revealed by the existence of divergence in beta.

1. Introduction

The European Union (EU) fosters international agreement and cooperation to promote smart, sustainable and inclusive growth of Member States. At the same time, the EU acknowledges Member States' national interests and sovereignty, allowing for cultural features and specificity of sectors in different countries. In the case of the education and training sector, depending on the country, the education strategy is formulated at the local, regional or national level, ensuring alignment with the strategic framework for European cooperation in education and training. A careful monitoring of education systems performance can be a decisive factor for EU to aspire a leading role in the education market worldwide.

Educational effectiveness or educational quality, which are frequently used as synonyms, can be defined as the degree to which an education system achieves the desired goals and effects (Burušić, Babarović, and Velić, 2016). Education systems, with their available resources (typically limited), should strive to maximise the outcomes attained.

Albeit EU countries are responsible for their education systems, EU policy aims to support national actions. Since 2009, the European Commission monitors the performance of education systems in Member States according to the strategic framework "Education and Training 2020" (ET2020) (Council of the European Union, 2009). This European cooperation framework has set four common objectives to address challenges in education and training for all Member States by 2020. These are as follows: (i) Making lifelong learning and mobility a reality; (ii) Improving the quality and efficiency of education and training; (iii) Promoting equity, social cohesion, and active citizenship; (iv) Enhancing creativity and innovation, including entrepreneurship, at all levels of education and training. To achieve these objectives, a set of targets have also been stated. These common objectives and targets seek convergence in terms of the achievements envisaged for European countries' education systems.

Our study is focused on performance benchmarking of European educational systems. It evaluates the initial relative position of the European countries' educational systems and explores the evolution of performance in the light of ET2020 targets. It uses the most recent data currently available in databases that sup-

* Corresponding author.
E-mail addresses: acamanho@fe.up.pt (A.S. Camanho), dovile.stumbriene@mif.vu.lt (D. Stumbriene), flavia.barbosa@inesctec.pt (F. Barbosa), audrone.jakaitiene@mf.vu.lt (A. Jakaitiene).

https://doi.org/10.1016/j.ejor.2022.05.048

port international comparisons. A Global Malmquist Index, specified with a metafrontier, is used to obtain a summary measure of the evolution of performance in the period 2009–2018. The index is estimated for triennial periods (coinciding with PISA waves) within this time window to understand trends in more detail. We also analyse the root sources of performance change, i.e., efficiency change and best practice change (frontier shift) for each triennial period. This is complemented with an analysis of sigma and beta convergence to obtain an overall picture of movements towards the best-practice frontier (sigma-convergence), as well as the evolution of the gap between the best and worst practice frontiers (beta-convergence).

Our contribution to the literature is twofold: from a methodological perspective, we propose a new approach for estimating sigma and beta convergence in a BoD setting. This required the formulation of a specific linear programming model to estimate a BoD worst practice frontier in the presence of weight restrictions. We also used a Global Malmquist Index, estimated with a metafrontier, that enables exploring the evolution of performance over time, satisfying the property of circularity in comparisons involving several periods. From an empirical perspective, we explored the trends in the performance of European Education Systems during the ET2020 Education and Training strategic framework. This is a very relevant topic given the importance of this sector for the development of nations and prosperity of societies.

This research extends previous research dedicated to the estimation of convergence using non-parametric frontier models that was initiated in Horta & Camanho (2015) for DEA evaluations and followed by Pereira, Camanho, Marques, and Figueira (2021) for a Directional BoD context.

The remainder of this paper is organised as follows. Section 2 reviews the studies focusing on the evaluation of education systems performance. Section 3 presents a Global Malmquist Index adapted to a BoD context, as well as its components of Efficiency Change and Best Practice Change. It also develops the methodological approach that enables the estimation of sigma and beta-convergence using BoD models with weight restrictions. Section 4 presents the indicators used and the data collected. Section 5 discusses the findings regarding the evolution of education and training performance in EU countries in the period 2009–2018. Section 6 concludes the paper, outlining the methodological contributions as well as the main findings and future research directions.

2. Literature review on performance assessment in the education sector

The empirical study reported in this paper is focused on performance benchmarking of European educational systems and explores the trends of the last decade. Accordingly, we reviewed the literature on the evaluation of education system performance, with a particular focus on analysis conducted at the country level, as well as studies exploring the evolution of performance over time.

2.1. Efficiency measurement in education

The assessment of performance of educational systems is a common practice in many countries. This requires the collection of outcome indicators for different educational stages and resources/funding available. Frontier methods enable obtaining an overall picture of efficiency and effectiveness of educational systems.

Educational systems have been widely studied at various levels. A detailed literature review by De Witte & López-Torres (2017) concluded that most studies focused either at the higher education level or basic/secondary school level. Secondary education studies are usually concerned with the ability of schools to promote the achievement of good academic results for their students, given the initial ability levels and socio-economic context. Higher education studies are typically concerned with the efficient use of resources. These assessments explore institutions' ability to minimise costs or resources used in the production of outputs (students education, research production, or third mission activities).

Education studies have used data at different levels of aggregation, ranging from analysis at the student, classroom or school level, to regional analysis at city, district or country level. In higher education studies, the levels considered have been the university department, faculty, university or research infrastructures. From a methodological perspective, frontier techniques or multilevel regression are typically applied in this scientific field. Frontier techniques enable comparisons to identify best-practices and guide improvements. Multilevel regression studies aim to evaluate the impact of different factors, organised according to a hierarchical structure, on students' performance.

2.2. Cross-country comparisons

Only a few studies were focused on international benchmarking comparisons, based on data collected at national level. Given that comparable national datasets, such as PISA, TIMSS or PIRLS, are increasingly available, the number of cross-country comparison is likely to rise soon.

Table 1 provides an overview of the papers available in the literature focusing on cross-country comparisons of performance using frontier techniques, considering the country as the unit of assessment.

Table 1 shows that most studies (10 out of 16) focused exclusively on a single educational stage (9 papers on secondary education and one paper on tertiary education). However, comparisons based only on secondary or tertiary education stages do not represent the performance of the overall educational system. A few studies (6) combined several educational stages, providing a more comprehensive perspective of the performance of the educational system. Clements (2002) explored compulsory education, comprising both primary and secondary education. Hanushek & Luque (2003), Agasisti, Munda, and Hippe (2019) and Ahec-Šonje, Deskar-Škrbić, and Šonje (2018) covered secondary and tertiary educational stages. Stumbriene, Camanho, and Jakaitiene (2019) considered the overall outcomes of European educational systems, including indicators related to primary, secondary and tertiary education, as well as adult participation in lifelong learning. Bogetoft, Heinesen, and Tranæs (2015) analysed education production (considering all educational levels) in Nordic countries in comparison with other OECD countries. The indicators covered both quantitative issues (the number of students enrolled in educational programs) as well as quality measures of educational production (graduation/completion rates and expected earnings after completion of studies).

The majority of papers (14 out of 16) used data from large-scale international assessments (10 articles used PISA and 4 papers used TIMSS). A few studies (7) combined data from PISA or TIMSS waves with data from other international databases (OECD, UNESCO or World Bank). Only the studies of Agasisti (2011) and Bogetoft et al. (2015) analysed the performance of education systems without using data from large-scale international assessments. Agasisti (2011) measured the efficiency of higher education systems in European countries using data from the OECD database (accessibility of the system, availability of financial resources and human resources as inputs, and graduation rates, employability of graduates and attractiveness of the systems as outputs). The data used in Bogetoft et al. (2015) came from OECD and Eurostat databases. The input was expenditure per student, and the outputs were the num-

Table 1
Studies on educational performance at country level.

Study	Stage			Data			Method						
							parametric		non-parametric			change over time	
	P	S	T	PISA	TIMSS	Other	SFA	Reg	FDH	DDF	DEA	MI	MLI
Afonso & Aubyn (2005)		x		x		x			x		x		
Afonso & Aubyn (2006)		x		x		x		x			x		
Agasisti (2011)			x			x					x		
Agasisti (2014)		x		x		x		x			x	x	
Agasisti et al. (2019)	x	x		x		x					x		
Bogetoft et al. (2015)	x	x	x			x					x		
Clements (2002)	x	x			x	x			x				
Giambona, Vassallo, and Vassiliadis (2011)		x		x							x		
Giménez, Prior, and Thieme (2007)		x			x						x		
Giménez et al. (2017)		x		x					x		x	x	
Giménez et al. (2019)		x			x					x	x		x
Hanushek & Luque (2003)	x	x			x			x					
Ahec-Šonje et al. (2018)		x	x	x		x					x		
Stumbriene et al. (2019)	x	x	x	x		x					x		
Sutherland, Price, and Gonand (2010)		x		x			x						
Thieme et al. (2012)		x		x						x			

P: primary education, S: secondary education, T: tertiary education MI: Malmquist Index, MLI: Malmquist-Luenberger Index.

ber of students enrolled, graduation rates, completion rates, and expected earnings.

As shown in Table 1, Data Envelopment Analysis (DEA) is the most widely used method to assess efficiency in education at country level (12 out of 16 papers). Only a few studies used other methods, such as Free Disposable Hull (FDH, 3 papers), regression (3 papers), Directional Distance Function (DDF, 2 papers) or Stochastic Frontier Analysis (SFA, 1 paper). In Hanushek and Luque (2003), the production function was estimated using regression, so it does not represent the best possible performance evaluated against a frontier, reflecting only average country-level performance.

In Agasisti et al. (2019), the DEA analysis was complemented with Multiple Criteria Decision Analysis to test the robustness of the DEA results. The authors proposed finding “compromise solutions” combining DEA and MCDA, as an efficient solution is not necessarily preferred to every non-efficient solutions. This study illustrates the importance of corroborating the results of traditional frontier based methods using non-frontier based mathematical approaches, to support decision making in real-world settings and align decisions with stakeholders preferences.

The study of Thieme, Giménez, and Prior (2012) was the first to specify desirable outputs (PISA scores) and undesirable outputs (inequality of educational achievements) in education efficiency studies at country level. The results obtained showed that it is feasible to combine high achievement with low inequality, and this is the ideal target that should be pursued in all countries.

There is another strand of the literature that explores performance at sub-national level. For example, Camanho, Varriale, Barbosa, and Sobral (2021) compared the performance of upper secondary schools (*Liceo*) among Italian Macro-regions, and Le, Afsharian, and Ahn (2021) compared the efficiency of the educational system in provinces across Vietnam. The goal was to reveal how well a province transforms the family expenditure in education into students’ achievements, suggesting targets that can improve the education system.

2.3. Assessment of performance trends over time

Most studies of education systems efficiency involving international comparisons focused only on one point in time. Only the papers of Agasisti (2014), Giménez, Thieme, Prior, and Tortosa-Ausina (2017) and Giménez, Thieme, Prior, and Tortosa-Ausina (2019) analysed performance change over time at country level. Table 2 provides detailed information about these studies. Note, however, that assessments of the evolution of performance over time are common in within country analysis of specific educational stages (e.g., Agasisti, Shibanova, Platonova, and Lisyutkin, 2020; Agasisti, Yang, Song, and Tran, 2021; Aparicio, Ortiz, and Santín, 2021; López-Torres & Prior, 2020).

The time periods, countries analysed and methodological approaches differ among the studies presented in Table 2. The outputs of all studies are based on PISA or TIMMS international assessments. Agasisti (2014) used PISA scores as outputs, while the other two studies (Giménez et al., 2017; 2019) considered both good (or desirable) outputs representing students’ achievements as well as bad (or undesirable) outputs reflecting educational inequality.

The results of Agasisti (2014) confirmed that a convergence process was ongoing in European countries in the period 2006–2009, but this process was quite slow (efficiency increased from 0.89 to 0.91 and standard deviation decreased from 0.042 to 0.034). The main results of Giménez et al. (2019) and Giménez et al. (2017) were aligned. The global index revealed productivity improvements in European educational systems, with a positive value of efficiency change, and a slight drop in technological change (or frontier shift). Giménez et al. (2017) identified three groups of European countries in terms of trends in productivity change, based on the results of the Global Malmquist Index and its components. The first group (Austria, Czech Republic, Germany, Italy, Latvia, Luxembourg, Poland, Portugal and Spain) improved productivity due to efficiency improvements. The second group (Finland and Ireland) showed a stable profile in terms of productivity levels. The third group (Greece, Hungary, Slovak Republic and Sweden) worsened productivity due to unfavorable technological change.

Most studies reviewed (reported in Tables 1 and 2) are focused on education efficiency or productivity change, where countries are regarded as consuming a set of inputs that is transformed into a set of outputs. In this process, efficient units are those that use the least inputs (e.g., expenditure, learning hours, teaching quality, student-teacher ratio) to maximises good outputs (e.g., TIMSS or PISA scores) while simultaneously minimising bad outputs (e.g., standard deviation of scores

Table 2
studies on educational performance at country level over time.

Study	Periods	Countries	Inputs	Outputs	Methodology
Agasisti (2014)	2006, 2009	20 European countries	Student-teacher ratio, expenditure per student	PISA scores	DEA with bootstrapping, regression, Malmquist Index
Giménez et al. (2017)	2003, 2012	29 countries (16 European countries)	Ratio of teachers per 100 students, index of quality of physical infrastructures, index of economic, social and cultural status	PISA scores, standard deviation of PISA scores	DEA, FDH, global non-radial Malmquist Index
Giménez et al. (2019)	2007, 2011	28 countries (9 European countries)	Learning hours, teaching quality	TIMSS scores, proportion of students with results below the basic learning standards	DEA, DDF, Global Malmquist Luenberger Index

or proportion of students with results below the basic learning standards - if considered in the assessment). This is the most popular approach followed in the education benchmarking literature.

2.4. Assessment of education achievements using BoD composite indicators

There is an alternative perspective of performance assessment, which is exclusively focused on the evaluation of achievements (outputs) without taking into account resource constraints. This approach was first adopted for the evaluation of educational systems at the country level in Stumbriene et al. (2019), which resorted to the estimation of Benefit-of-the-Doubt (BoD) composite indicators. The fundamental difference between these two approaches (DEA efficiency versus BoD performance evaluations) relates to whether or not resource constraints are taken into account in the attainment of outputs. Stumbriene et al. (2019) proposed a composite indicator to summarise the overall performance of education systems in the light of Europe 2020 strategy. The DEA model included weight restrictions that were customised to range from fixed weights to fully flexible weights. This approach enabled testing the robustness of the performance score to varying degrees of flexibility in the weights, reflecting different perspectives of decision-makers about the importance of exhibiting well-balanced performance in all dimensions of educational achievements.

Composite indicators are often used to aggregate data on the performance of individual processes to obtain a simplified view of complex systems. In the literature, only a few studies used this approach for performance evaluations in the education context. De Witte & Hudrlikova (2013) applied a robust conditional BoD model to rank universities, allowing for flexibility in the weighting system to account for different profiles in terms of universities' strengths. This method also accounts for background characteristics of universities and evaluates which characteristics have a stronger impact on the ranking. De Witte & Rogge (2010) used a robust conditional BoD model to assess the research output of faculty members at a large polytechnic university in Belgium. The robust formulation allows mitigating the effect of atypical observations in the sample, as well as reducing the impact of measurement errors. The model also specified weight restrictions based on stakeholder opinions, to enhance the acceptance of the results. The conditional formulation allowed taking into account background conditions affecting researchers' productivity (e.g., motivation, gender, employment conditions). De Witte & Schiltz (2018) used a robust conditional BoD composite indicator to measure organisational effectiveness at the school district level. The analysis was based on data from surveys and interviews to school board members and principals in Flanders, the northern part of Belgium.

As far as we know, there are no studies where the composite indicator (BoD approach) is used to evaluate the evolution of performance over time in the education context, although we can find a few applications in other fields.

Rogge (2019) constructed a geometric Benefit-of-the-Doubt composite indicator to measure EU Member States' progress towards the achievement of Europe 2020 strategy (EU2020), which covers five policy domains: employment, research and development, climate change and energy sustainability, education, and social inclusion and poverty reduction. For the analysis of the education domain, the author only used two ET2020 framework indicators, namely 'Early Leavers from Education and Training' and 'Tertiary Educational Attainment'.

This paper contributes to the benchmarking literature in the education sector by evaluating the evolution of performance in European countries' education systems in the light of ET2020 targets. It extends the cross-section evaluation conducted in Stumbriene et al. (2019), based on the BoD approach. It also builds on the study of Rogge (2019) by focusing in more detail on the evolution of the educational domain of the strategic framework ET2020. Our research focuses exclusively on the educational sector, using a more comprehensive set of indicators and a longer time period to assess educational systems' performance trends.

3. Methodology

This section presents the BoD framework, which underlies the evaluation of education systems performance conducted in this study. It also explains the Global Malmquist Index satisfying the circular property that is estimated using a metatechnology. Its components of Efficiency Change and Best Practice Change are also discussed. Finally, the concepts of sigma and beta convergence are described, and the framework used for their estimation in a BoD context with weight restrictions is developed.

3.1. The BoD setting

The analysis of the performance of European educational systems described in this study involves the estimation of a composite indicator, according to the Benefit-of-the-Doubt (BoD) framework proposed by Cherchye, Moesen, Rogge, and Puyenbroeck (2007). The composite indicator aggregates several individual indicators into a summary measure of performance that provides an overall perspective of educational achievements at the country level.

The BoD formulation is equivalent to an original DEA model (Charnes, Cooper, and Rhodes, 1978), assuming constant returns to scale, with a dummy input equal to one. Following Koopmans (1951), this dummy input can be interpreted as an 'helmsman' attempting to steer the country towards better performance.

We also incorporate in the model information regarding the relative importance of the indicators in the form of weight restrictions, following Zanella, Camanho, and Dias (2015). These restrictions are expressed in the form of Assurance Regions Type I (ARI), which impose bounds on the ratios between the weights (Thompson, Langemeier, Lee, Lee, and Thrall, 1990). These are the most common type of weight restrictions used in DEA assessments. The formulation of the restrictions as proposed in Zanella et al. (2015) has the advantage of allowing the specification of bounds in percentage terms, reflecting the relative importance of the outputs considered in the assessment.

Consider a DMU k under assessment, observed in period t $(t = 1, \ldots, P)$ with the vector of output indicators $Y_k^t = (y_{1k}^t, y_{2k}^t, \ldots, y_{sk}^t)$. The performance of this DMU can be compared with other DMUs j $(j = 1, \ldots n)$ of the sample, either observed in the same period t or in any other period in the sample. The frontier underlying the comparison envelops the production possibility set (or technology set, denoted by T^t) of the time period t considered in the assessment.

The formulation of the BoD model with the weight restrictions is shown in (1). Note that the incorporation of restrictions on output weights requires the use of an output oriented formulation of the BoD model, as in Stumbriene et al. (2019). Although BoD formulation is usually presented with an input orientation, Van Puyenbroeck (2018) discussed the direct, reciprocal relation between the BoD-model and the output-oriented DEA model introduced by Lovell, Pastor, and Turner (1995). According to Van Puyenbroeck (2018), it is more intuitive to solve an input-oriented model when the problem is exclusively concerned with the aggregation of outputs. However, as we also wish to incorporate value judgements on the relative importance of the output indicators, this requires the use of an output oriented formulation of the BoD model.

$$\begin{aligned}
\left[E_k^t(\mathbb{1}, Y_k^t)\right]^{-1} = \text{Min } & \nu \\
\text{s.t. } & \sum_{r=1}^{s} u_r y_{rk}^t = 1 \\
& \sum_{r=1}^{s} u_r y_{rj}^t - \nu \le 0 \quad j = 1, \ldots, n \\
& \frac{u_r \bar{y}_r}{\sum_{r=1}^{s} u_r \bar{y}_r} \ge \phi_r \quad r = 1, \ldots, s \\
& u_r \ge 0 \qquad r = 1, \ldots, s \\
& \nu \ge 0
\end{aligned} \tag{1}$$

The decision variables u_r, $r = 1, \ldots s$, and ν correspond to weights associated with the output indicators and the unitary dummy input, respectively. The restriction of output weights corresponds to the third restriction in model (1), where ϕ_r is the percentage of the total virtual output associated with output r, for an artificial DMU whose outputs correspond to the average value of the outputs r (r=1,... s) observed for all DMUs in the sample in all time periods, e.g. ($\bar{Y} = (\bar{y_1}, \bar{y_2}, \ldots, \bar{y_s})$).

We imposed weight restrictions on the minimum importance given to each output indicator, to ensure that all education dimensions are taken into account in the performance assessment. The weight restrictions in model (1) ensure that all countries must allocate at least ϕ_r% of the total weight ($\phi_r = 0.05$) to each indicator r $(r = 1, \ldots, s)$. They also enable enhancing the discriminatory power of the benchmarking model.

The composite indicator measure $E_k^t(\mathbb{1}, Y_k^t)$ is obtained by comparison with the best practice frontier of period t, and is given by the multiplicative inverse of the objective function value of model (1).

The deterministic nature of DEA models means that all the deviations of observed units from the frontier are exclusively attributed to inefficiency. A well-know limitation of deterministic frontier estimation techniques is that the existence of outliers in the sample may lead to the shift of the frontier, and significantly impact the evaluation of several other units. In this context, detecting observations that differ to a large extent from the rest of the sample is critical. The literature proposes different approaches to mitigate the effect of outliers, which include procedures to detect and remove influential observations in DEA (e.g. Wilson, 1995), or robust order-m methods based on statistical bootstrapping that allow using partial frontiers instead of full frontiers to reduce the impact of outliers in the assessment (Aragon, Daouia, and Thomas-Agnan, 2005; Cazals, Florens, and Simar, 2002). In this paper we adopted a traditional approach of outliers detection prior to the estimation of efficiency. As no significant outliers were identified, we proceeded with the use of a full frontier model in the empirical analysis reported in this paper.

3.2. Global Malmquist Index

To explore the evolution of performance over time, in this paper, we use a formulation of the Malmquist Index that satisfies the circular property, which requires defining a metatechnology that envelops all data.

The concept of a metafrontier function was first introduced by Battese & Rao (2002). They considered a metafrontier to be an envelope function of the frontiers of different group technologies, estimated by stochastic frontiers. This approach was enhanced by Battese, Rao, and O'Donnell (2004) that defined the metafrontier as a deterministic non-parametric production function. It assumes each period p has its own technology T^p $(p = 1, 2, \ldots, t, t+1, \ldots, P)$, but there is a metatechnology $T^M = \bigcup_{t=1}^{P} T^t$ which envelops all individual period technologies.

Afsharian & Ahn (2015) developed an overall Malmquist index based on a non-convex metatechnology, to overcome the non-homogeneity problem related to the determination of the global benchmark technology when measuring productivity change over time. However, the estimation of efficiency using the model proposed by Afsharian & Ahn (2015) is computational demanding in the presence of large samples evaluated in several time periods. Therefore, in this paper, we adopt the convex estimation of the Global Malmquist Index proposed by Pastor & Lovell (2005). Based on the concept of a metatechnology, Pastor & Lovell (2005) proposed a Global Malmquist Productivity Index with a technology formed from data of all units in all periods. This index overcomes the limitation of non-circularity of the original Malmquist Index of Färe, Grosskopf, Lindgren, and Roos (1989, 1992). Such Global Malmquist Index (GMI) is defined regarding a unique metatechnology and provides a cross-sectional comparison of the performance of all periods. The specification of the index for a multi-input multi-output setting, as proposed in Pastor & Lovell (2005), is shown in (2).

$$GMI_k^{t,t+1} = \frac{E_k^M\left(X_k^{t+1}, Y_k^{t+1}\right)}{E_k^M(X_k^t, Y_k^t)} \tag{2}$$

This index can be adapted to the context of a BoD evaluation as shown in (3):

$$GMI_k^{t,t+1} = \frac{E_k^M\left(\mathbb{1}, Y_k^{t+1}\right)}{E_k^M(\mathbb{1}, Y_k^t)} \tag{3}$$

The BoD formulation considering a metatechnology can be obtained by adapting the DEA linear programming formulation proposed by O'Donnell, Rao, and Battese (2008). The resulting BoD formulation, also with the inclusion of weight restrictions, is shown in (4).

$$\begin{aligned}
\left[E_k^M(\mathbb{1}, Y_k^t)\right]^{-1} = \text{Min } & \nu \\
\text{s.t. } & \sum_{r=1}^{s} u_r y_{rk}^t = 1 \\
& \sum_{r=1}^{s} u_r y_{rj}^t - \nu \le 0 \; j = 1, \ldots, n, \; t = 1, \ldots, P
\end{aligned}$$

$$\frac{u_r\bar{y}_r}{\sum_{r=1}^{s} u_r\bar{y}_r} \geq \phi_r \quad r = 1, \ldots, s$$
$$u_r \geq 0 \qquad r = 1, \ldots, s$$
$$v \geq 0 \tag{4}$$

The composite indicator measure $E_k^M(\mathbb{1}, Y_k^t)$ estimates the performance of DMU k from period t considering a metatechnology, and is given by the inverse of the objective function of model (4). The generalisation to the assessment of performance of a DMU k in time period $t+1$, corresponding to the component $E_k^M(\mathbb{1}, Y_k^{t+1})$, is straightforward.

This index can also be decomposed in the Efficiency Change component and Best-Practice Change component, representing the frontier shift, as follows:

$$\begin{aligned} GMI_k^{t,t+1} &= \frac{E_k^{t+1}(\mathbb{1}, Y_k^{t+1})}{E_k^t(\mathbb{1}, Y_k^t)} \times \left(\frac{E_k^M(\mathbb{1}, Y_k^{t+1})}{E_k^{t+1}(\mathbb{1}, Y_k^{t+1})} \times \frac{E_k^t(\mathbb{1}, Y_k^t)}{E_k^M(\mathbb{1}, Y_k^t)} \right) \\ &= \frac{E_k^{t+1}(\mathbb{1}, Y_k^{t+1})}{E_k^t(\mathbb{1}, Y_k^t)} \times \left(\frac{\frac{E_k^M(\mathbb{1}, Y_k^{t+1})}{E_k^{t+1}(\mathbb{1}, Y_k^{t+1})}}{\frac{E_k^M(\mathbb{1}, Y_k^t)}{E_k^t(\mathbb{1}, Y_k^t)}} \right) \\ &= EC_k^{t,t+1} \times \left(\frac{BPG_k^{t+1,M}(\mathbb{1}, Y_k^{t+1})}{BPG_k^{t,M}(\mathbb{1}, Y_k^t)} \right) \\ &= EC_k^{t,t+1} \times BPC_k^{t,t+1} \end{aligned} \tag{5}$$

$EC_k^{t,t+1}$ is the efficiency change component that assesses the evolution of the distance to the frontier between period t and $t+1$, for a DMU k under assessment in each of the these periods. $EC_k^{t,t+1} \geq 1$ indicates greater proximity to the frontier in $t+1$ than in t.

$BPG_k^{t,M}(\mathbb{1}, Y_k^t)$ is the best practice gap between the frontier in period t and the metafrontier. This gap is measured along the ray representing the output mix of the DMU k under assessment in period t. The interpretation of $BPG_k^{t+1,M}(\mathbb{1}, Y_k^{t+1})$ is equivalent, but for time period $t+1$.

$BPC_k^{t,t+1}$ is the change in BPG between t and $t+1$, and provides an estimate of frontier shift between t and $t+1$. $BPC_k^{t,t+1} \geq 1$ indicates that the benchmark technology in period $t+1$ (for the output mix observed in $t+1$ for DMU k) is closer to the global benchmark technology (the metatechnology) than it was in period t (for the output mix observed in t for DMU k).

To improve the discriminatory power, the Malmquist Index is estimated with weight restrictions added to the local technologies T^t and metatechnology T^M, which is aligned with the index originally proposed by Alirezaee & Afsharian (2010).

The previous definition of the Global Malmquist Index can be used to make comparisons involving more than two time periods satisfying the circular relationship, as follows:

$$GMI_k^{t,t+2} = GMI_k^{t,t+1} \times GMI_k^{t+1,t+2} \tag{6}$$

This relationship indicates that the index estimating the evolution of performance between t to $t+2$ is equal to the product of the index comparing t to $t+1$ and the index comparing $t+1$ to $t+2$. The same property can be observed for the components EC and BPC.

3.3. The assessment of convergence

This research explores the estimation of sigma and beta convergence in a BoD setting. The σ-convergence analyses if the dispersion of the DMU's distance to the within-period frontier tends to decrease over time. The β-convergence analyses the evolution of the distance between the best and worst-practice frontier over time. This requires the estimation of the movements of the DMUs located on the best practice frontier and also the movements of the DMUs located on the worst practice frontier.

3.3.1. σ-Convergence

To determine σ-convergence, we follow the approach proposed by Horta & Camanho (2015) which consists of calculating the geometric mean of $EC_j^{t,t+1}$ for all DMUs in the sample, as shown in (7):

$$\sigma\text{-convergence} = \left(\prod_{j=1}^{n} EC_j^{t,t+1} \right)^{\frac{1}{n}} \tag{7}$$

If σ-convergence > 1, then there is convergence (i.e. the DMUs moved closer to the best practice frontier from period t to $t+1$). σ-convergence < 1 means divergence (i.e. the DMUs moved away from the best practice frontier between t to $t+1$). σ-convergence $= 1$ indicates that, on average, the DMUs are located at a similar distance to the frontier in period t and $t+1$.

3.3.2. β-Convergence

To evaluate β-convergence, we need to analyse the movements of the DMUs located on the best practice metafrontier and the movements of the DMUs located on the worst practice metafrontier. Note that the best and worst metafrontiers are estimated on the basis of the same metatechnology T^M, corresponding to the union of the local technologies T^t. However, these frontiers correspond to the envelopment of data from top and bottom performance levels, leading to the concepts of a 'best metafrontier' and a 'worst metafrontier'.

We follow the approach proposed by Horta & Camanho (2015) to estimate β-convergence, which consists of calculating the ratio of the Best Practice Change (BPC) to the Worst Practice Change (WPC) for the DMUs in the sample, and then computing the geometric average of these values, as shown in (8):

$$\beta\text{-convergence} = \left(\prod_{j=1}^{n} \frac{BPC_j^{t,t+1}}{WPC_j^{t,t+1}} \right)^{\frac{1}{n}} \tag{8}$$

Best Practice Change ($BPC_j^{t,t+1}$) is computed as described in (5). Worst practice change is estimated using expression (9), adopting a similar procedure to the estimation of Best Practice Change discussed in the previous section. This requires the estimation of composite indicator values $W_k^t(\mathbb{1}, Y_k^t)$ and $W_k^t(\mathbb{1}, Y_k^{t+1})$ measured against the worst practice frontier of period t and $t+1$, respectively, as well as estimates of performance compared with the metafrontier: $W_k^M(\mathbb{1}, Y_k^t)$ and $W_k^M(\mathbb{1}, Y_k^{t+1})$.

$$WPC_k^{t,t+1} = \left(\frac{\frac{W_k^M(\mathbb{1}, Y_k^{t+1})}{W_k^{t+1}(\mathbb{1}, Y_k^{t+1})}}{\frac{W_k^M(\mathbb{1}, Y_k^t)}{W_k^t(\mathbb{1}, Y_k^t)}} \right) = \left(\frac{WPG_k^{t+1,M}(\mathbb{1}, Y_k^{t+1})}{WPG_k^{t,M}(\mathbb{1}, Y_k^t)} \right) \tag{9}$$

We propose in (10) an inverted BoD model, in which the distances are measured in relation to the worst practice frontier. The model shown in (10) enables the estimation of $W^t(\mathbb{1}, Y^t)$ incorporating weight restrictions.

$$\left[W_k^t(\mathbb{1}, Y_k^t)\right]^{-1} = \text{Max} \quad v$$
$$\text{s.t.} \quad \sum_{r=1}^{s} u_r y_{rk}^t = 1$$
$$v - \sum_{r=1}^{s} u_r y_{rj}^t \leq 0 \quad j = 1, \ldots, n$$
$$\frac{u_r\bar{y}_r}{\sum_{r=1}^{s} u_r\bar{y}_r} \geq \phi_r \quad r = 1, \ldots, s$$
$$u_r \geq 0 \qquad r = 1, \ldots, s$$
$$v \geq 0 \tag{10}$$

Table 3
Description of ET2020 indicators used in the analysis.

Indicator	Description	Source
Y_1	**Early leavers from education and training**. The percentage of the population aged 18–24 with at most lower secondary education and who were not in further education or training during the last four weeks	Eurostat
Y_2	**Tertiary educational attainment**. The share of population aged 30–34 years who have successfully completed university or university-like (tertiary-level) education	Eurostat
Y_3	**Early childhood education and care**. The share of population aged four to the age when the compulsory education starts who are participating in early education	Eurostat
Y_4	**Employment rates of recent graduates**. The share of employed graduates (20–34 years) having left education and training 1–3 years before the reference year	Eurostat
Y_5	**Adult participation in learning**. The participation rate of adults (25–64 years) in education and training (last 4 weeks)	Eurostat
Y_6	**Underachievement in reading, mathematics and science**. The average percentage of PISA Low achievers (below Level 2) in reading, mathematics and science	OECD

The assessment of performance of DMU k in period t evaluated against the worst practice metafrontier $W^M(\mathbb{1}, Y^t)$ is given by the inverse of the objective function of model (10) if, in the second constraint, we replace the set of DMUs observed in period t (corresponding to $j = 1, \ldots, n$) by the set $j = 1, \ldots, n$ for all periods under assessment $t = 1, \ldots, P$ that comprise the metatechnology.

If β-convergence > 1, then the worst and the best practice frontiers are more distant in time period $t+1$ than in period t (i.e., there is divergence between periods t and $t+1$). If β-convergence < 1, then both frontiers are closer in period $t+1$ than in period t (i.e., there is convergence between periods t and $t+1$).

4. Indicators and sample selected

As described in the introduction, the European Commission uses the Education and Training Monitor initiative to foster performance enhancements in the education systems of Member Countries. The European Commission selected eight key indicators for monitoring, and benchmark values to be reached by 2020 in the context of ET2020.

The empirical analysis reported in the next section uses six of the key indicators defined. We could not use the mobility indicators because 'Learning Mobility in Higher Education' only has partial data available, and 'Learning Mobility in Initial Vocational Education and Training' has no data available.

The description of the key indicators used in this study is provided in Table 3. Note that some indicators are focused on past performance, such as Tertiary Education Attainment (Y2) or Employment Rates of Recent Graduates (Y4). In contrast, others measure actual performance, such as Early Leavers from Education and Training (Y1) and Adult Participation in Learning (Y5). As we monitor the achievement of ET2020 targets at the country level, all indicators are considered in the performance assessment, irrespectively of the time frame covered.

All data were extracted from Eurostat and OECD databases for 31 European countries for the years 2009, 2012, 2015 and 2018 (European Commission, 2021, visited on 26th of August 2021). The European Commission monitors the performance of 27 European Union countries (EU-27). In this study, we also included four European countries not in the European Union (UK, Iceland, Norway, and Switzerland) to make the analysis more comprehensive. Our final sample comprises 31 countries.

Missing data at country level occurred for variables Y_2, Y_4 and Y_6. For 2009, we do not have information about Y_2 values for Cyprus and Malta. Therefore we filled this gap with the most recent year with data available (2012). Y_4 was missing in 2009 for Norway, and the value was replaced with data of 2012. Y_6 was missing in 2009 for Cyprus (2012 data replaced the missing values), 2012 for Malta (2015 data replaced the missing values) and 2009 for Austria (2012 data replaced the missing values). PISA 2018 data for Spain only has values for two domains (math and science), as reading information is missing. Therefore the average of Y_6 for Spain in 2018 is calculated only with the values of math and science.

For the calculation of the BoD scores, the indicators Y_1 (early leavers from education and training) and Y_6 (underachievement in reading, maths and science) were adjusted by subtracting the values observed in each country from 100%. This ensures that higher values of all indicators analysed correspond to better performance.

Table 4 provides the main descriptive statistics for the analysed variables, from which we can explore performance trends for each individual indicator. It also reports the goals proposed by the ET2020 strategic framework for each indicator for countries in the EU-27 region. The goals can be summarised as follows: The rates of early school leavers (Y_1) should be reduced to below 10%. At least 40% of 30–34-year-olds should have completed tertiary education (Y_2). At least 95% of children between the age of four and the age for starting compulsory (ISCED 1) should participate in education (Y_3). Note that the age for starting compulsory education is different from country to country (European Commission, 2017).

Concerning employment rates of recent graduates (Y_4) the goal is to reach values above 82% by 2020. Recent graduates refers to the employment rates of persons aged 20 to 34 fulfilling the following conditions: being employed according to the International Labour organisation (ILO) definition; having attained at least upper secondary education (ISCED 3) as the highest level of education; not having received any education or training in the four weeks preceding the survey; having successfully completed their highest educational attainment 1, 2 or 3 years before the survey. The rate of adult participation in learning (Y_5) should reach at least 15%. Finally, the share of 15 years-old with underachievement in reading, mathematics and science (Y_6) should be less than 15% across EU Member states (underachievers means failing Level 2 on the PISA scale for mathematics, reading and science).

We see convergence towards the benchmark value for early childhood education and care (Y_3), as the average values are approaching the target over time with smaller variance. The average value of adult participation in learning (Y_5) improved over the four periods considered, but the standard deviation does not diminish, indicating that convergence of countries' performance is hardly happening. The average share of low achievers in basic skills (Y_6) does not show signs of convergence towards the benchmark, indicating a large scope for improvement in this indicator for most European countries. Concerning the indicators Y_1, Y_2 and Y_4, the average value for the European Countries studied shows that the targets were exceeded, although a few countries failed to meet the goal.

Fig. 1 visually illustrates the performance of each country concerning the attainment of ET2020 goals two years before the milestone (year 2018). This Figure shows for each country the number

Table 4
Descriptive statistics of ET2020 indicators for 31 European countries.

Variable	Statistics	2009	2012	2015	2018	ET2020 target
Y_1	Average	13.0	11.3	10.0	9.5	below 10
	Standard Deviation	6.8	5.3	4.6	4.5	
	Minimum	4.9	4.4	2.8	3.3	
	25-percentile	8.7	7.5	6.9	6.2	
	75-percentile	14.7	12.5	11.6	11.3	
	Maximum	30.9	24.7	20.2	21.5	
	No. countries above target	20	17	13	12	
Y_2	Average	34.2	37.4	41.4	43.7	above 40
	SD	10.6	9.7	9.1	9.0	
	Min	16.8	21.7	25.3	24.6	
	25-percentile	24.0	27.8	32.3	34.9	
	75-percentile	42.0	43.9	47.6	49.4	
	Max	50.4	52.2	57.6	57.6	
	No. countries below target	16	16	11	10	
Y_3	Average	88.9	91.0	92.4	92.7	above 95
	SD	9.6	8.6	6.9	7.1	
	Min	69.2	71.7	73.8	73.6	
	25-percentile	84.2	85.5	89.2	91.9	
	75-percentile	96.1	97.4	97.6	97.4	
	Max	100.0	100.0	100.0	100.0	
	No. countries below target	21	16	15	14	
Y_4	Average	79.4	76.1	77.5	82.8	above 82
	SD	7.6	11.2	11.4	9.1	
	Min	60.6	43.0	45.2	55.3	
	25-percentile	74.4	70.2	72.4	80.6	
	75-percentile	84.5	84.1	84.7	88.1	
	Max	92.9	92.5	95.0	94.8	
	No. countries below target	19	20	19	11	
Y_5	Average	10.9	11.7	12.3	12.8	above 15
	SD	8.1	8.7	9.1	8.4	
	Min	1.6	1.4	1.3	0.9	
	25-percentile	4.7	5.4	5.8	6.6	
	75-percentile	17.4	16.3	18.6	19.1	
	Max	31.4	31.6	31.5	31.6	
	No. countries below target	23	23	21	20	
Y_6	Average	21.2	20.8	22.5	23.4	below 15
	SD	8.1	7.9	7.9	8.4	
	Min	7.3	8.2	10.2	10.0	
	25-percentile	17.2	16.1	16.9	18.7	
	75-percentile	20.9	22.0	25.1	24.3	
	Max	42.9	40.0	40.5	46.0	
	No. countries above target	27	24	27	27	

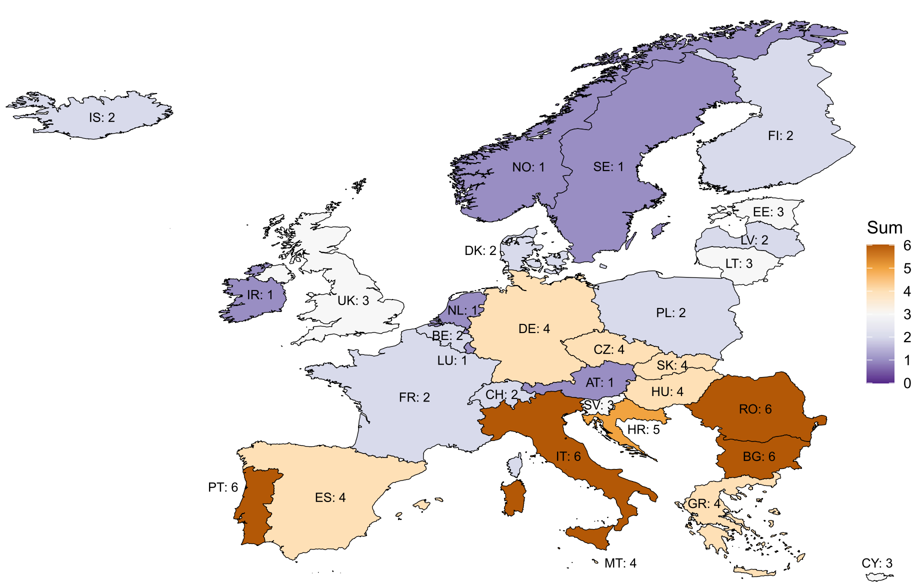

Fig. 1. Number of indicators (out of 6) that missed the ET2020 target in 2018.

of indicators that failed the ET2020 target in 2018. In Fig. 1 dark orange colors reflect worse performance. We observed quite an heterogeneous context in terms of the achievement of targets among EU countries.

Before proceeding to the estimation of performance using BoD models, we explored the existence of outliers in the sample studied. Outliers are a reason of concern for most empirical methods, particularly so in DEA, because an outlier may shift the position of the frontier and significantly affect the efficiency score of the units under assessment. These observations, which may result from measurement errors, are sometimes called high leverage points and may be eliminated such that the model is not distorted to fit these extreme points. In a benchmarking setting, the units located on the frontier should correspond to replicable performance, so exceptionally high relative performance units may be removed from the sample for precautionary reasons. Thus, attempting to ensure that the frontier is well populated and not too distant from the bulk of DMUs in the analysis is an important step of any efficiency assessment.

Considering that outliers are observations located below the first quartile minus three times the interquartile range for indicators Y_1 and Y_6 (with lower values representing better performance), or observations located above the third quartile plus three times the interquartile range for indicators Y_2 to Y_5, we can conclude that the sample studied does not have outliers. We could only identify marginal outlier values, with a value lower than the first quartile minus 1.5 times the interquartile range, for indicator Y_6 in 2009 (for Estonia and Finland) and in 2018 (for Estonia). We verified that these values do cause distortions to the location of the frontier, so we proceeded with the estimation of efficiency using a deterministic BoD model based on all data available in the sample, without removal of marginal outliers.

5. Results regarding the evolution of education and training in the period 2009–2018

This section discusses the initial relative position of countries' performance in the light of ET2020 Education and Training targets. The discussion follows with a characterisation of the evolution of performance at the country level in the period 2009–2018, and finalises with an overview of sigma and beta convergence for European countries educational systems achievements.

5.1. BoD results

First, we report the results of the BoD composite indicator (with weight restrictions), revealing the relative position of countries performance in each of the four years analysed. This assessment considers as reference the best practice frontier observed in each year. In this context, the BoD score provides a single summary measure of the distance of each country to the Best Practice frontier at the period considered.

The BoD performance score reported in Table 5 aims to quantify the distance between each country's current situation (in terms of the indicators for which the European Commission has defined a common goal), and the best practices observed in other EU countries. Note that the BoD model aggregates the values of all indicators of education achievements without taking into account the available resources, so it does not intend to provide a measure of efficiency in terms of value for the money invested.

Table 5 shows that the countries occupying the last positions of the relative performance ranking in the first year analysed (2009) are Romania and Bulgaria. These countries remained at the bottom of the ranking until 2018. Greece occupied one of the lowest positions of the rank until 2015, but improved its relative position in 2018.

The only country that remained on the best practice frontier in all years analysed is Sweden. The other countries that have been at the frontier in three of the years considered are Switzerland, Iceland, Finland, Denmark and Norway. However, two different cases should be distinguished: Switzerland and Iceland reveal an improvement trend (these two countries were at the frontier in all years studied except for the first year considered - 2009). Conversely, Denmark and Norway moved away from the frontier in the last year considered (both countries were efficient in all years except 2018).

In the most recent year under study (2018), two countries exhibit a notable improvements in the rank compared with the previous period: Estonia and Luxembourg. Estonia improved all the six ET2020 indicators considered in the last year analysed, compared to 2015. Since 2014/15, there have been substantial changes in early childhood education and care and graduate employability. Since 2014, support for parents of children in early childhood education and care has been expanded, and Estonia started using labour market forecasting systematically in 2016 (European Commission, EACEA & Eurydice, 2020). Furthermore, the Estonian government adopted a lifelong learning strategy, including combating early school leaving in 2014, as well as introduced a reform in higher education based on improving quality and changing the funding system in 2013 (European Commission, 2015). The Estonian school system combines excellence in teaching basic skills with a high degree of equity. The reduction of early leavers from education and training (especially the dropout in upper secondary education) remains a key challenge for the country (European Commission, 2020a).

Luxembourg has made progress in improving tertiary education attainment, employment rates of recent graduates and early leavers from education and training in the period 2015–2018. Luxembourg has set itself the target of raising the tertiary attainment rate among 30–34 year-olds to 66% by 2020. In 2019, the rate was 56.2% and is the third highest in the EU. Furthermore, the employment rate of recent tertiary graduates in 2019 was 94.2% (EU average: 85.0%), reflecting strong demand for highly skilled workers (European Commission, 2020e). Since 2016/2017, the orientation process at the end of primary school has been reformed (European Commission, 2018b). Pupils' performance is heavily influenced by their ability to cope with the trilingual system. Luxembourg introduced a multilingualism education programme targeting children aged one to four (European Commission, EACEA & Eurydice, 2020) as well as provided language support in the migrant students' mother tongue (European Commission, EACEA & Eurydice, 2016) in recent years. The reduction of a large proportion of underachievers in a highly diverse school population is a key challenge for the country (European Commission, 2020e).

Note that being on the frontier means that there is no scope for equiproportional improvement in all indicators simultaneously in the year considered, based on the evidences observed in other European countries. However, as we are estimating performance with a radial model, non-radial improvements may still exist for some countries located on the frontier.

5.2. Results of the Global Malmquist Index

The results of the Malmquist Index provide an overview of the evolution of Education and Training among EU countries. The first column of Table 6 shows the value of the MI for the period 2009/2018, representing a summary measure of the performance change observed in each country in the decade analysed.

The results reveal a trend of performance improvement for most countries in the period 2009–2018, which is in line with the results obtained by Rogge (2019) for the period 2008–2014. On average, the change in performance of EU countries in this period

Table 5
DEA results for all years analysed $E^t(\mathbb{1}, Y^t)$.

Country	2009		2012		2015		2018	
(j)	$E_j^t(\mathbb{1}, Y_j^t)$	rank	$E_j^t(\mathbb{1}, Y_j^t)$	rank	$E_j^t(\mathbb{1}, Y_j^t)$	rank	$E_j^t(\mathbb{1}, Y_j^t)$	rank
Austria	0.932	17	0.966	12	0.937	15	0.933	17
Belgium	0.956	13	0.912	21	0.901	19	0.926	20
Bulgaria	0.818	30	0.813	29	0.810	30	0.802	30
Croatia	0.893	25	0.862	27	0.875	24	0.873	26
Cyprus	0.992	10	0.946	14	0.957	13	0.962	14
Czechia	0.921	20	0.916	19	0.893	22	0.909	22
Denmark	1	1	1	1	1	1	0.995	8
Estonia	0.948	15	0.977	10	0.963	12	1	1
Finland	1	1	1	1	0.990	7	1	1
France	0.954	14	0.913	20	0.953	14	0.969	12
Germany	0.923	19	0.939	16	0.933	16	0.927	18
Greece	0.827	29	0.813	30	0.839	29	0.863	27
Hungary	0.883	26	0.848	28	0.871	25	0.880	24
Iceland	0.994	9	1	1	1	1	1	1
Ireland	0.998	7	1	1	0.978	9	0.999	7
Italy	0.908	22	0.871	24	0.848	28	0.857	28
Latvia	0.873	27	0.892	23	0.879	23	0.897	23
Lithuania	0.902	23	0.931	17	1	1	0.974	11
Luxembourg	0.996	8	0.985	8	0.986	8	1	1
Malta	0.929	18	0.945	15	0.928	18	0.945	15
Netherlands	1	1	0.983	9	0.977	10	0.984	9
Norway	1	1	1	1	1	1	0.977	10
Poland	0.920	21	0.918	18	0.899	20	0.927	19
Portugal	0.853	28	0.868	25	0.861	26	0.879	25
Romania	0.805	31	0.770	31	0.765	31	0.766	31
Slovakia	0.897	24	0.865	26	0.849	27	0.852	29
Slovenia	0.966	12	0.950	13	0.930	17	0.936	16
Spain	0.944	16	0.899	22	0.894	21	0.913	21
Sweden	1	1	1	1	1	1	1	1
Switzerland	0.999	6	1	1	1	1	1	1
UK	0.990	11	0.970	11	0.964	11	0.962	13
Average	0.936		0.927		0.925		0.933	
Std Deviation	0.058		0.064		0.063		0.062	
Min	0.805		0.770		0.765		0.766	
Max	1		1		1		1	
25-percentile	0.900		0.882		0.877		0.889	
Median	0.944		0.939		0.933		0.936	
75-percentile	0.995		0.984		0.982		0.990	

was 4.4%, as shown in the summary statistics of Table 6. Only one country (Slovakia) shows clear evidence of decline in performance, with a magnitude of 4%.

Regarding the analysis of the components of the GMI for each triennial period, we can see that performance improvements are due to a combined effect of gains in productivity at the frontier and relative efficiency improvements. The frontier shift component signaling gains at the frontier shows values greater than one in 2009–2012 (BPC = 1.031) and 2012–2015 (BPC = 1.018), as shown in Table 6. In addition, countries moved closer to the annual European best practice frontier in the most recent years (as shown by $EC = 1.008$ for the period 2015–2018), meaning that education and training achievements are becoming more homogeneous in the EU space.

The countries with the greatest performance improvement in the decade analysed are Portugal, Spain, Malta and Greece, all of them with a magnitude of performance change greater than 10%. In the case of Portugal and Spain, the improvement is remarkable, with a magnitude of 28% and 19%, respectively, mostly explained by the progress observed in the first triennial period 2009–2012).

Specifically, Greece and Portugal occupied low positions in the performance ranking in 2009 (Greece was in the 29th position and Portugal in the 28th position). Given their low starting point, it means that they succeed in the effort of moving towards the best practices observed in peer EU countries. In the case of Greece, the improvement has been evenly balanced among the triennial periods studied, whilst Portugal showed greater improvement in the beginning of the period studied, and a stabilisation in the most recent years. The case of Spain and Malta are notable, in the sense that despite being in the middle of the ranking in the first year studied (2009), they also managed to improve significantly in the last decade.

Next, the countries with the greatest performance improvement are discussed, and the overview of reforms and policy developments are presented. Portugal moved from a very low position and has the greatest performance improvement in the decade analysed. Portugal has significantly reduced its early school leaving rate, and tertiary education attainment has significantly improved in the decade analysed. Since 2012, Portugal has implemented a national programme to address school failure and early school leaving, with a new monitoring system introduced in 2013/14. Many new programmes and measures were introduced over the last decade to ensure equity in basic education. Still, the high proportion of students re-sitting years and the extent to which socioeconomic background determines educational achievement demonstrates that equity is a crucial challenge for the country (European Commission, 2015). A number of measures have been taken to ease and widen enrolment to higher education (reducing fees and increasing the number of available university places, scholarships, and student housing facilities) in the decade analysed. Higher education enrollment and tertiary attainment kept rising in the decade analysed, but only 30% of students who enter a bachelor programme grad-

Table 6
Evolution of performance achievements of European countries in the period 2009–2018.

	2009/2018				2009/2012			2012/2015			2015/2018		
Country (j)	GMI_j	EC_j	BPC_j	Rank	GMI_j	EC_j	BPC_j	GMI_j	EC_j	BPC_j	GMI_j	EC_j	BPC_j
Austria	1.016	1.001	1.016	20	1.011	1.036	0.975	1.005	0.970	1.037	1	0.995	1.005
Belgium	1.028	0.968	1.062	15	0.990	0.954	1.038	1.022	0.989	1.033	1.017	1.027	0.990
Bulgaria	1.023	0.981	1.043	17	1.026	0.995	1.031	0.990	0.995	0.994	1.008	0.991	1.017
Croatia	1.020	0.978	1.043	18	1.001	0.965	1.038	1.024	1.016	1.008	0.995	0.998	0.997
Cyprus	1.044	0.970	1.077	13	1.003	0.953	1.053	1.070	1.012	1.057	0.973	1.006	0.967
Czechia	0.992	0.987	1.004	29	0.999	0.995	1.004	0.993	0.975	1.018	1	1.018	0.983
Denmark	1.012	0.995	1.018	22	1.021	1	1.021	1.017	1	1.017	0.975	0.995	0.980
Estonia	1.025	1.055	0.972	16	1.037	1.030	1.007	0.979	0.986	0.992	1.010	1.038	0.973
Finland	1.018	1	1.018	19	1.011	1	1.011	0.997	0.990	1.007	1.010	1.010	0.999
France	1.042	1.015	1.027	14	1.007	0.956	1.053	1.029	1.045	0.985	1.006	1.016	0.990
Germany	1.009	1.004	1.005	24	1.007	1.017	0.990	1.004	0.994	1.011	0.998	0.994	1.004
Greece	1.111	1.044	1.064	4	1.034	0.983	1.052	1.038	1.032	1.006	1.035	1.029	1.006
Hungary	0.989	0.996	0.992	30	0.997	0.960	1.038	1.002	1.027	0.976	0.990	1.010	0.980
Iceland	0.997	1.006	0.991	27	1.015	1.006	1.009	1.016	1	1.016	0.967	1	0.967
Ireland	1.077	1.001	1.076	6	1.022	1.002	1.019	1.034	0.978	1.057	1.019	1.021	0.999
Italy	1.057	0.943	1.120	9	1.022	0.959	1.066	1.031	0.974	1.059	1.002	1.010	0.992
Latvia	1.070	1.027	1.042	7	1.043	1.021	1.022	1.008	0.986	1.022	1.018	1.020	0.998
Lithuania	1.045	1.080	0.968	12	1.024	1.032	0.992	1.011	1.075	0.940	1.010	0.974	1.037
Luxembourg	1.015	1.004	1.011	21	0.996	0.989	1.007	0.987	1.001	0.986	1.033	1.014	1.018
Malta	1.112	1.017	1.093	3	1.054	1.017	1.036	1.019	0.982	1.038	1.035	1.018	1.017
Netherlands	1.045	0.984	1.062	11	1.027	0.983	1.045	1.008	0.994	1.014	1.010	1.007	1.003
Norway	1.093	0.977	1.119	5	1.034	1	1.034	1.054	1	1.054	1.003	0.977	1.027
Poland	1.005	1.008	0.998	25	0.996	0.998	0.998	1.004	0.979	1.025	1.005	1.031	0.975
Portugal	1.276	1.030	1.239	1	1.151	1.017	1.131	1.086	0.992	1.094	1.022	1.021	1.001
Romania	1.002	0.952	1.053	26	0.986	0.957	1.030	0.984	0.993	0.991	1.033	1.001	1.032
Slovakia	0.961	0.950	1.011	31	0.996	0.964	1.033	0.983	0.982	1.002	0.982	1.004	0.978
Slovenia	1.012	0.969	1.044	23	1.010	0.983	1.027	0.994	0.980	1.014	1.008	1.007	1.002
Spain	1.188	0.967	1.228	2	1.090	0.952	1.144	1.062	0.994	1.069	1.026	1.022	1.004
Sweden	0.995	1	0.995	28	0.995	1	0.995	1.005	1	1.005	0.995	1	0.995
Switzerland	1.052	1.001	1.050	10	1.040	1.001	1.038	1.003	1	1.003	1.008	1	1.008
UK	1.059	0.972	1.090	8	1.027	0.979	1.049	1.030	0.994	1.036	1.001	0.999	1.002
Average	1.044	0.997	1.048		1.022	0.991	1.031	1.015	0.998	1.018	1.006	1.008	0.998
Std Deviation	0.062	0.030	0.063		0.032	0.025	0.035	0.026	0.022	0.031	0.017	0.015	0.018
Min	0.961	0.943	0.968		0.986	0.952	0.975	0.979	0.970	0.940	0.967	0.974	0.967
Max	1.276	1.080	1.239		1.151	1.036	1.144	1.086	1.075	1.094	1.035	1.038	1.037
25-percentile	1.010	0.977	1.007		1	0.964	1.008	0.998	0.983	1.002	0.999	1	0.985
Median	1.024	1	1.043		1.013	0.997	1.031	1.008	0.994	1.014	1.008	1.009	0.999
75-percentile	1.056	1.008	1.064		1.032	1.005	1.038	1.028	1	1.036	1.018	1.020	1.006

uate within three years (the expected duration of the programme for most fields) (European Commission, 2020g). Performance-based funding mechanisms with a focus on returning to study have been provided to higher education institutions (European Commission, EACEA & Eurydice, 2015).

The second largest performance improvement has been made by Spain. Spain has made remarkable progress in reducing its early school leaving rate in the decade analysed. Despite a steady fall in early school leaving over the decade, Spain still has the second highest rate in Europe, with significant differences between regions. Since 2014/15, Spain has implemented the law for improvement of the quality of education based on improving student performance and curbing early school leaving. The law has been implemented at different paths across the different autonomous communities. Furthermore, to make a vocational education and training path more attractive as an alternative to leaving school early, Spain has introduced two-year initial vocational training programs for students aged 15 (European Commission, 2015). In the past five years, reforms to the reinforcement of existing support related to student background have been taken (European Commission, EACEA & Eurydice, 2020).

The other two countries with a performance change greater than 10% are Malta and Greece. Malta has significantly reduced its early school leaving rate as well as improved the tertiary education attainment and life-long learning rates. Malta has been investing significantly in its education and training system in the decade analysed. Since 2014, Malta has implemented the early school leaving strategy and the literacy strategy. Both strategies acknowledge the need to improve professional development of teachers at all stages of their career (European Commission, 2015). Since 2014, Malta has introduced reforms to ensure that education and career guidance are provided not only through school-based guidance, but also through the national curriculum (at primary and secondary school), thus systematically reaching all students. In the past five years, reforms in targeted funds have led to the establishment of a scheme for additional support to disadvantaged students (European Commission, EACEA & Eurydice, 2020). Malta has created a quality assurance framework for further and higher education (European Commission, 2015). Tertiary educational attainment has increased further, mainly due to the arrival of EU nationals on the labour market. Despite the very high employment rate of recent tertiary graduates (95% vs. 85% at EU level in 2019), skills shortages remain an issue at all levels (European Commission, 2020f).

The Greek education and training sector had several important structural reforms that took place under the economic adjustment programme, which lasted until 2015. The Greek education system was significantly rationalised at primary and secondary key stages between 2011 and 2014 (European Commission, 2015). Greece has upgraded the vocational education and training

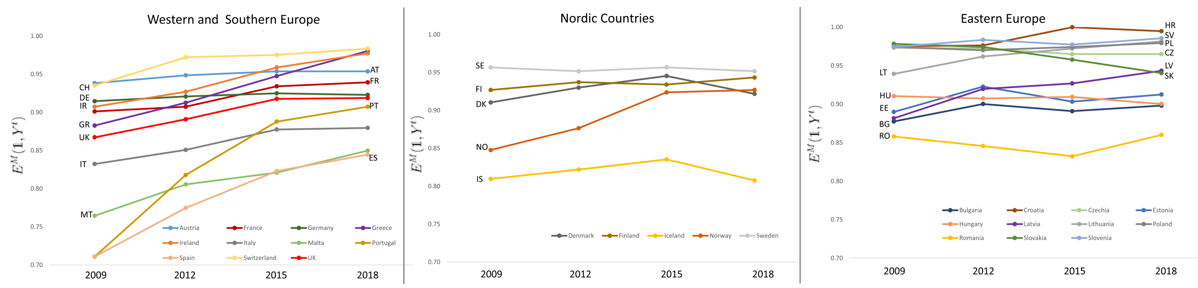

Fig. 2. Evolution of performance in the period 2009–2018.

sector - the vocational lyceum was reformed to allow for greater permeability between programmes and promote a smoother transition from one education pathway to another in 2016 (European Commission, EACEA & Eurydice, 2020). In 2013/14, Greece carried out a comprehensive reform of the governance and organisation of its higher education system (European Commission, 2015). Furthermore, Greece started using labour market forecasting systematically in 2015 (European Commission, EACEA & Eurydice, 2020). Although the employment of tertiary graduates has risen, their lack of soft skills (communication, teamwork, self-confidence, and work ethic) still affects their job prospects (European Commission, 2020b).

Fig. 2 graphically illustrates the trends for selected countries. It reports the values of the Composite Indicator computed against the Metafrontier $E^M(\mathbb{1}, Y^t)$, so it can be interpreted as a performance trend, where lower values represent lower productivity compared with the best values of the sample observed in the decade analysed. The results of the performance evaluation against the metafrontier underlying the construction of the graphs in Fig. 2 are provided on Table A.1 in Appendix A.

Regarding western and southern European Countries, most countries show a strictly improvement trend in all periods considered. As discussed previously, Portugal, Spain, and Malta show evidence of closing the gap in relation to the best performance levels observed in this region. Italy also started in a relatively low rank position, but despite evidence of improvement between 2009 and 2018, it is still below the best performing countries. Ireland and Switzerland are notable cases in the sense that they occupy the top positions of the performance rank in 2018, with consistent improvement trends in the decade analysed.

Ireland has one of the largest improvement trends among western and southern European Countries and occupies the top position of the performance rank in 2018. In the decade analysed, Ireland has undertook reforms to increase quality, relevance and achievement at every level of its education system. Ireland's national strategy to improve literacy and numeracy among children and young people has been carried out during the period 2011–2020 (European Commission, 2015). Irish government launched a comprehensive action plan for education (2016–2019) which contains a set of actions to be implemented with particular focus on disadvantage, skills, and continuous improvement within the education sector (its goal is to make Irish education one of the best performers in Europe by 2026) (European Commission, 2018a). The focus on equity in education has made Irish secondary schools positive forces for inclusion and social mobility. The reduced proportion of low achieving students and comparatively low variation between schools reflect the effectiveness of the measures taken in recent years to create an equitable and high-performing education system (European Commission, 2020c). A structural reform agenda has been set out for the Irish higher education sector in 2011, witch sets out a new vision for higher education in Ireland (European Commission, 2015). Ireland has a very high tertiary attainment rate (enrolment in higher education has traditionally been high in Ireland, reflecting at least in part the lack of alternatives and the relative undervaluation of vocational pathways) (European Commission, 2018a). Between 2016 and 2020, significant modernisation and expansion in apprenticeships has been implemented (European Commission, 2020c).

Analysing the Nordic Countries, it is noticeable the improvement trend of Norway between 2009 and 2015, that converged towards the best practice values observed in Sweden, Finland, and Denmark. Denmark also showed an improvement trend between 2009 and 2015, but declined in the most recent period (2018). Iceland shows a similar trend to Denmark, but its performance levels are considerably lower than those of the other Nordic countries.

Norway has significantly reduced its early school leaving rate, as well as the increase in tertiary education attainment and employment rates of recent graduates in the period 2009–2018. The participation rate for children aged 4 and older in early childhood education and care continued to be high. Since 2015/16, Norway has invested heavily in early childhood education and care and taken measures to remove financial barriers to early childhood education and care and increase participation among underrepresented groups (low-income families and families of immigrant backgrounds). Since 2016, the alternative vocational education and training pathway has been adopted (the scheme was piloted between 2008 and 2011), which improves completion rates among students who are struggling in mainstream vocational education and training. Furthermore, Norway developed and implemented the national qualifications framework for lifelong learning (2011) and the Norwegian strategy for skills policy (2017–2021) to define the qualifications on different levels (from the end of lower secondary education to doctoral level) as well as to ensure the skills that enhance the labour market relevance of graduates. In 2017, Norway revised the performance-based component of the higher education funding model to include indicators on completion rates and to reward institutions that secure third-party funding (OECD, 2020).

Regarding Eastern European Countries, the largest performance improvements occurred for Latvia (7%) and Lithuania (4.5%). The case of Croatia is also worth noting, as it was among the best performing countries at the beginning of the period analysed and also improved marginally (2%) in the period studied, so in 2018 it occupies the top performance rank among Eastern European countries (followed by Slovenia, Lithuania and the Czech Republic). In 2018, Slovakia, and Latvia exhibit similar levels of performance, but reached that position with opposite trends, as Slovakia's performance declined in the period 2012–2018. Estonia, Hungary, and Bulgaria occupy low positions in the ranking of Eastern European Education systems' performance. Romania occupied the last posi-

Table 7
Convergence results.

time period	σ-convergence	β-convergence	average BPC	average WPC
2009-2018	0.996	1.055	1.045	0.991
2009-2012	0.991	1.032	1.030	0.998
2012-2015	0.998	1.016	1.016	1.000
2015-2018	1.008	1.006	0.999	0.993

tion in all years analysed, with a declining trend between 2009–2015, although it was able to recover in the last period (2015–2018). Thus, this country was unable to converge towards the performance levels observed in peer countries, lagging considerably below the EU2020 targets.

In the decade analysed, Latvia has made remarkable progress in reducing its early school leaving rate, increasing its tertiary education attainment, employment rate of recent graduates, and attendance in early childhood education and care. Since 2016, Latvia has guaranteed a free place in early childhood education from the age of 18 months for all children whose parents so wish (since 2013, Latvia has provided financial support for parents whose children aged 18 months to four years do not have a place in public early childhood education) (European Commission, 2020d). To identify causes of early school leaving and thus improve the evidence base for policy actions, a number of studies were carried out in 2014 and 2015 (European Commission, 2015). The early school leaving rate is well below the EU average but rises significantly outside cities, echoing geographical disparities in learning outcomes. Latvian tertiary graduates have the highest employment rate in the EU, pointing to significant returns to higher education. The employment rate of recent tertiary graduates is significantly higher than the EU average of 85%, while the employment rate of vocational education and training graduates is one of the lowest in the EU (European Commission, 2020d).

5.3. Analysis of convergence

Although EU countries design their own education systems, they develop common political goals and work closely together at EU level. This provides impetus for national developments of education system as well as convergence towards the EU targets.

Table 7 summarises the values of convergence estimated for the period considered (for the decade 2009–2018, as well as the detailed results for triennial periods). The estimation of β-convergence requires the estimation of performance against the best and the worst practice frontiers. The results obtained for these estimates, obtained using models (1) and (10), are reported on Table B.1 in Appendix B. Note that a value of $W_k^t(\mathbb{1}, Y_k^t) = 1$ indicates that the country is located on the WPF on year t, and values lower than 1 indicate that the country is above the WPF.

The results of σ-convergence are obtained using expression (7), and the results of β-convergence are obtained using expression (8). The components of β-convergence (BPC) are (WPC) obtained using expressions (5) and (9), respectively.

Recall that values of σ-convergence > 1 are desirable, meaning that, on average, the DMUs of the sample moved closer to the Best-Practice Frontier over time. The only evidence of convergence in σ is observed in the last period considered (2015–2018), as revealed by the value of σ-convergence equal to 1,008. Many of policy interventions and reforms at the national level since the establishment of the strategic framework Education and Training 2020, have focused on achieving the benchmark values of key indicators. The results obtained indicate that European countries moved closer to the European best practice frontier especially in the last period analysed (2015–2018).

This movement of approximation to the frontier can be evaluated at country level by the value of the efficiency change component of the Malmquist Index. The evolution in terms of proximity to the frontier for each country in the period 2009–2018 (EC component of the GMI) is illustrated in Fig. 3. It can observe significant improvements in Lithuania, Estonia, Greece and Portugal.

Values of β-convergence < 1 are desirable, meaning that on average the Best-Practice Frontier and Worst-Practice Frontiers moved closer together over time. There is no evidence of β-convergence in the years analysed. Although there is evidence of performance improvements in the Best-Practice Frontier in the first triennial periods (see the values of average BPC > 1 in Table 7), and stability in the last triennial period (average BPC ≈ 1 in the period 2015–2018), this happened alongside stability or decline in performance on countries located on the worst practice frontier in the periods studied (see the values of average WPC ≤ 1 on Table 7).

Concerning the trends in individual dimensions of the ET2020 strategy, it is observed convergence in some policy areas in the last decade. For example, six structural indicators were defined to monitor early leavers from education and training. As stated in the European Commission report (European Commission, EACEA & Eurydice, 2020) the majority of European countries accomplished policy interventions according to the defined indicators: 1) the majority of European countries are collecting national data on early leavers from education and training through a student register; 2) almost all European countries now have policies to promote alternative education and training pathways, and many countries also aim to facilitate transitions between the different pathways; 3) almost all European countries have policies for language support for students with a different mother tongue; 4) two-way approach to promoting education and career guidance in schools is promoted through top-level policies in about two-thirds of European countries; 5) almost all European countries have policies promoting second chance education for early leavers, and most of them support early leavers through targeted education and career guidance. The exception is a structural indicator of addressing early leaving from education and training in initial teacher education and continuing professional development – this is the area where the least number of policies can be found. Since 2015, inclusive education has become the focus of recommendations and practical support in the context of ET2020 (European Commission, EACEA & Eurydice, 2020). However, despite the emphasis on inclusive education, policy shortages in this area remain an issue in many countries.

Despite the observed convergence of the indicator representing early childhood education and care (the average value of the indicator approaches the target over time with a smaller variance), the quality of this educational stage differs according to children's age. Almost half of the European countries guarantee a place in early childhood education and care from age three (often free of charge). Furthermore, a growing number of countries make attendance compulsory during the last year of early childhood education and care. However, quality requirements are lower for working with younger children than older ones (European Commission, EACEA & Eurydice, 2019). For example, less than half of European countries have the minimum qualification requirement for staff working with children under three years old, and only two thirds of countries have guidelines for working with younger chil-

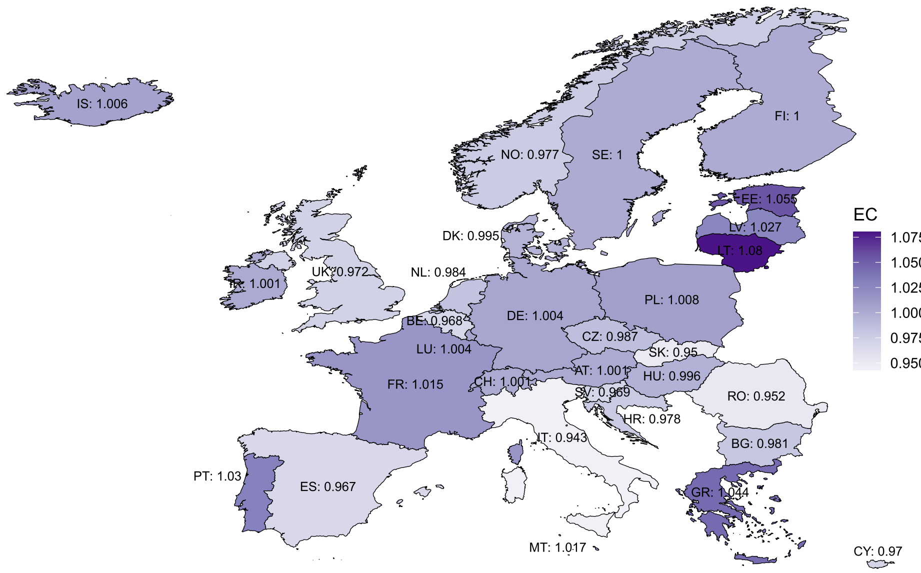

Fig. 3. Efficiency change.

dren (European Commission, EACEA & Eurydice, 2020). In contrast, educational guidelines for working with older children are established in all European countries, and almost all countries have the minimum qualification requirements for staff working with older children. (European Commission, EACEA & Eurydice, 2020).

Another example could be from the higher education area, where the convergence in the policy domain is ambiguous. Although structural indicators related to the quality and relevance of higher education (monitoring the characteristics of the student body, recognition of informal and non-formal learning, and completion rate as a requirement in external quality assurance) were implemented in the majority of education systems, indicators on equity (targets for widening participation of under-represented groups, and performance-based funding with focus on social dimension) are still waiting to progress in more than half European countries (European Commission, EACEA & Eurydice, 2020). The indicator that represents tertiary educational attainment in the ET2020 framework does not reflect the equity dimension. As stated in the report on the implementation of the strategic framework (Council of the European Union, 2015), the focus of the policy needs to be re-calibrated to include both the pressing economic and employment challenges and the role of education in promoting equity and inclusion.

6. Conclusions

The main contribution of this paper from a methodological perspective is the development of a new framework applicable in a BoD setting that allows the estimation of a Malmquist Index and the analysis of σ and β convergence. Although a few studies used the BoD approach to evaluate the evolution of performance over time, in the education context there are no studies evaluating performance trends using composite indicators. From an empirical perspective, the contribution of this paper is the assessment of the performance changes in European education systems in the period 2009–2018, focusing on the attainment of ET2020 targets. This study has assessed the performance of education systems in 31 European countries, providing new tools for monitoring progress and identifying critical challenges, and contributing to evidence-based policy making.

In the decade analysed, some countries (Portugal, Spain, Malta, Greece, Norway, Ireland, and Latvia) reveal considerable performance improvements (more than 7%). Overall, the improvements observed in EU countries are due to a combined effect of gains in productivity at the frontier and relative efficiency improvements (countries moving closer to the best practices observed in peers). This suggests that the reforms and policy measures that have been implemented to move closer to the ET2020 targets have been fruitful. However, some countries (Slovakia and Hungary) exhibited performance declines in the period 2009–2018.

A heterogeneous situation is still observed in terms of the achievement of ET2020 goals. In different countries, the number of indicators that did not meet the ET2020 target in 2018 ranges from 0 to 6. Even though Portugal, Italy, Bulgaria, and Romania had not reached any of the ET2020 target values in 2018, three of these countries (Portugal, Malta, and Italy) improved their performance by more than 5% in the decade analysed. It should be acknowledged that the ET2020 targets should not be considered as specific national targets. Instead, the ET2020 targets are average values for European countries' performance by 2020. Member States were invited to consider how and to what extent they could contribute to the achievement of the European targets through their national actions, in line with national priorities and taking into account the economic circumstances.

Although the European Commission has set the strategic framework for European cooperation in education and training at the EU level, the success of its implementation is situated at the national level. As concluded by Wüst & Rogge (2021), the European Commission should continue to focus on establishing effective institutions and good governance at the country level. Our research findings show a mismatch between the targeted and achieved value of indicators, which creates space for rethinking the ET2020 targets to ensure more sustainable trajectories of improvement at the country level, accounting for the specificities of national educational contexts.

In some cases, it is evident that the defined ET2020 targets would not be achieved at the national level until 2020. A few examples include tertiary educational attainment in Romania, participation in early childhood education in Croatia and Switzerland, employment rates of recent graduates in Greece and Italy, adult participation in learning, and underachievement in reading, mathematics, and science in Romania and Bulgaria. This discrepancy between the actual and the target value of indicators can hinder the implementation of the framework at the EU level. Special support could be provided for countries lagging behind to adjust their national policies and contribute to a faster convergence towards the EU goals.

Furthermore, the peculiarities of national education systems and the socioeconomic context of the countries should be taken into account in performance assessments to obtain an unbiased picture of the differences among EU countries. For example, the low participation rate in early childhood education in Finland can be explained by the specificity of the early childhood education and care system in Finland, which provides freedom of choice for parents concerning whether or not to use state educational services (Stumbriene, Zelvys, Zilinskas, Dukynaite, and Jakaitiene, 2022). Also, the socio-economic context of post-socialist countries can affect the current level of educational performance. As stated by Želvys, Stumbriene, and Jakaitiene (2018), concepts like accountability, result-oriented management, or performance-based assessment still have different meanings in the East and the West. Furthermore, the shortage of trust and the abundance of control are noticeable in post-socialist countries' education systems. When the West is used as a single yardstick to measure post-socialist education transformations, it is not surprising that different patterns can be identified for eastern countries (Silova, Millei, Chachkhiani, Palandjian, and Vitrukh, 2021). Consequently, the EU targets may raise unreasonable expectations for the results of policy actions, as the progress of educational systems seems to occur in different leagues across EU countries.

Although convergence is the declared objective of the EU and is considered the mechanism to achieve European cohesion, the empirical findings of the current study found evidence of sigma-convergence only in the last period analysed (2015–2018). There are still a few countries lagging, such that the gap between the best and worst practice frontiers has increased, as revealed by the beta-convergence score larger than one for the period studied. The methodology proposed in this study opens new possibilities to assess convergence across countries (or regions), reflecting the evolution of several indicators using a single summary measure.

In its most general form, convergence is defined as reducing or equalizing disparities. In the EU context, the reduction of inequalities in society should be achieved among countries and regions. The ET2020 strategic framework highlights that the development of education must aim to help every citizen realise their full potential and to create sustainable economic prosperity in Europe[1]. However, the dimension 'every citizen' is not measurable by any of the selected indicators of the strategic framework. In the process of converting educational goals into performance indicators, the goals are reduced to several measurable indicators that do not cover the declared goal. As indicators become a key reference for education policy in EU countries, they are supposed to fully gauge the declared goal (including equity in society also). The results of this study indicate that a critical challenge for EU countries remains to ensure equity in education and training across the EU space.

Some limitations inevitably exist in the present study, which can be topics for further research. The methodology proposed in this paper can be extended to include a robustness assessment of trends, complementing the nonparametric techniques used in this paper with statistical methods, such as robust order-m approaches. Also, in this study, student achievement (used as an indicator of educational attainment) is taken from the PISA survey, which is conducted every three years. Thus, the performance estimation cannot be performed more frequently than three years due to the use of PISA indicators. However, triennial research can be limited in shaping national education policies. Historical data might not provide up-to-date analysis in a rapidly changing world. We hope this study will encourage researchers to analyse education performance over time at the national level more frequently, as this type of studies are still relatively rare. Furthermore, the present study deals only with outcome indicators, but it would be interesting to extend this assessment to an efficiency analysis, considering the resources used to attain the outcomes.

This study concludes by pointing out that, in February 2021, the Council agreed on a new strategic framework and set new EU-level targets for monitoring progress. This resulted in the approval by the Council of a Resolution on a strategic framework for European cooperation in education and training towards the European Education Area and beyond (2021–2030). This resolution outlines five strategic priorities for the period 2021–2030 and will monitor seven EU-level targets. The methodology proposed in this paper can be used as a tool to assess European countries' performance in the future.

Acknowledgements

This research is funded by the European Social Fund according to the activity "Improvement of researchers" qualification by implementing world-class R&D projects' of Measure No.09.3.3-LMT-K-712. The project No. DOTSUT-39 (09.3.3-LMT-K-712-01-0018) / LSS-250000-57.

Appendix A. Results of the performance evaluation against the metafrontier

Table A.1
DEA results for all years analysed against the metafrontier.

Country (j)	2009 $E_j^M(\mathbb{1}, Y_j^t)$	2012 $E_j^M(\mathbb{1}, Y_j^t)$	2015 $E_j^M(\mathbb{1}, Y_j^t)$	2018 $E_j^M(\mathbb{1}, Y_j^t)$
Austria	0.938	0.949	0.954	0.954
Belgium	0.915	0.905	0.925	0.940
Bulgaria	0.878	0.900	0.891	0.898
Croatia	0.975	0.976	1	0.995
Cyprus	0.908	0.912	0.975	0.949
Czechia	0.973	0.972	0.965	0.965
Denmark	0.910	0.930	0.945	0.922
Estonia	0.890	0.923	0.903	0.913
Finland	0.927	0.937	0.934	0.943
France	0.901	0.907	0.934	0.939
Germany	0.915	0.921	0.925	0.923
Greece	0.883	0.913	0.948	0.980
Hungary	0.910	0.907	0.909	0.900
Iceland	0.810	0.822	0.835	0.808
Ireland	0.907	0.927	0.959	0.977
Italy	0.832	0.851	0.878	0.880
Latvia	0.882	0.920	0.927	0.943
Lithuania	0.939	0.962	0.972	0.981
Luxembourg	0.950	0.945	0.933	0.964
Malta	0.764	0.806	0.821	0.850
Netherlands	0.913	0.937	0.944	0.954
Norway	0.848	0.877	0.924	0.927
Poland	0.974	0.970	0.974	0.979
Portugal	0.711	0.818	0.888	0.907
Romania	0.858	0.846	0.832	0.860
Slovakia	0.978	0.974	0.958	0.940
Slovenia	0.974	0.984	0.977	0.986
Spain	0.711	0.775	0.823	0.845
Sweden	0.957	0.952	0.957	0.952
Switzerland	0.935	0.972	0.975	0.984
UK	0.867	0.891	0.918	0.919

[1] https://eur-lex.europa.eu/legal-content/EN/LSU/?uri=celex:52009XG0528(01).

Appendix B. Results of the best and worst practice frontier

Table B.1
DEA results for all years analysed against the best and the worst practice frontier.

	2009		2012		2015		2018	
Country(j)	$E_j^t(\mathbb{1}, Y_j^t)$	$W_j^t(\mathbb{1}, Y_j^t)$	$E_j^t(\mathbb{1}, Y_j^t)$	$W_j^t(\mathbb{1}, Y_j^t)$	$E_j^t(\mathbb{1}, Y_j^t)$	$W_j^t(\mathbb{1}, Y_j^t)$	$E_j^t(\mathbb{1}, Y_j^t)$	$W_j^t(\mathbb{1}, Y_j^t)$
Austria	0.932	0.857	0.966	0.837	0.937	0.834	0.933	0.826
Belgium	0.956	0.862	0.912	0.871	0.901	0.853	0.926	0.854
Bulgaria	0.818	1	0.813	1	0.810	0.993	0.802	1
Croatia	0.893	1	0.862	1	0.875	1	0.873	0.981
Cyprus	0.992	0.910	0.946	0.929	0.957	0.925	0.962	0.885
Czechia	0.921	0.909	0.916	0.875	0.893	0.895	0.909	0.874
Denmark	1	0.785	1	0.772	1	0.762	0.995	0.809
Estonia	0.948	0.901	0.977	0.854	0.963	0.858	1	0.832
Finland	1	0.870	1	0.848	0.990	0.845	1	0.811
France	0.954	0.883	0.913	0.877	0.953	0.818	0.969	0.835
Germany	0.923	0.864	0.939	0.864	0.933	0.854	0.927	0.875
Greece	0.827	1	0.813	1	0.839	1	0.863	1
Hungary	0.883	0.932	0.848	0.913	0.871	0.881	0.880	0.907
Iceland	0.994	0.837	1	0.853	1	0.831	1	0.890
Ireland	0.998	0.932	1	0.855	0.978	0.831	0.999	0.809
Italy	0.908	1	0.871	1	0.848	1	0.857	1
Latvia	0.873	0.937	0.892	0.875	0.879	0.870	0.897	0.863
Lithuania	0.902	0.906	0.931	0.880	1	0.861	0.974	0.858
Luxembourg	0.996	0.823	0.985	0.813	0.986	0.809	1	0.802
Malta	0.929	0.991	0.945	0.973	0.928	0.946	0.945	0.926
Netherlands	1	0.804	0.983	0.810	0.977	0.800	0.984	0.802
Norway	1	0.829	1	0.836	1	0.796	0.977	0.816
Poland	0.920	0.956	0.918	0.888	0.899	0.885	0.927	0.860
Portugal	0.853	1	0.868	0.958	0.861	0.889	0.879	0.890
Romania	0.805	1	0.770	1	0.765	1	0.766	1
Slovakia	0.897	1	0.865	0.983	0.849	0.995	0.852	0.955
Slovenia	0.966	0.840	0.950	0.830	0.930	0.855	0.936	0.842
Spain	0.944	0.981	0.899	0.975	0.894	0.935	0.913	0.929
Sweden	1	0.796	1	0.775	1	0.774	1	0.766
Switzerland	0.999	0.836	1	0.811	1	0.820	1	0.856
UK	0.990	0.838	0.970	0.844	0.964	0.818	0.962	0.840

References

Afonso, A., & Aubyn, M. S. (2005). Non-parametric approaches to education and health efficiency in OECD countries. *Journal of Applied Economics, 8*(2), 227–246.

Afonso, A., & Aubyn, M. S. (2006). Cross-country efficiency of secondary education provision: A semi-parametric analysis with non-discretionary inputs. *Economic Modelling, 23*(3), 476–491.

Afsharian, M., & Ahn, H. (2015). The overall Malmquist index: A new approach for measuring productivity changes over time. *Annals of Operations Research, 226*(1), 1–27.

Agasisti, T. (2011). Performances and spending efficiency in higher education: A European comparison through non-parametric approaches. *Education Economics, 19*(2), 199–224.

Agasisti, T. (2014). The efficiency of public spending on education: An empirical comparison of EU countries. *European Journal of Education, 49*(4), 543–557.

Agasisti, T., Munda, G., & Hippe, R. (2019). Measuring the efficiency of European education systems by combining data envelopment analysis and multiple-criteria evaluation. *Journal of Productivity Analysis*, 1–20.

Agasisti, T., Shibanova, E., Platonova, D., & Lisyutkin, M. (2020). The Russian excellence initiative for higher education: A nonparametric evaluation of short-term results. *International Transactions in Operational Research, 27*(4), 1911–1929.

Agasisti, T., Yang, G.-l., Song, Y.-y., & Tran, C.-T. T. D. (2021). Evaluating the higher education productivity of Chinese and European “elite” universities using a meta-frontier approach. *Scientometrics*, 1–35.

Ahec-Šonje, A., Deskar-Škrbić, M., & Šonje, V. (2018). Efficiency of public expenditure on education: Comparing croatia with other NMS. In *INTED2018 Conference proceedings* (pp. 2317–2326).

Alirezaee, M., & Afsharian, M. (2010). Improving the discrimination of data envelopment analysis models in multiple time periods. *International Transactions in Operational Research, 17*(5), 667–679.

Aparicio, J., Ortiz, L., & Santín, D. (2021). Comparing group performance over time through the Luenberger productivity indicator: An application to school ownership in European countries. *European Journal of Operational Research*.

Aragon, Y., Daouia, A., & Thomas-Agnan, C. (2005). Nonparametric frontier estimation: A conditional quantile-based approach. *Econometric Theory, 21*(2), 358–389.

Battese, G. E., & Rao, D. P. (2002). Technology gap, efficiency, and a stochastic metafrontier function. *International Journal of Business and Economics, 1*(2), 87–93. https://EconPapers.repec.org/RePEc:ijb:journl:v:1:y:2002:i:2:p:87-93

Battese, G. E., Rao, D. S. P., & O'Donnell, C. J. (2004). A metafrontier production function for estimation of technical efficiencies and technology gaps for firms operating under different technologies. *Journal of Productivity Analysis, 21*(1), 91–103. https://doi.org/10.1023/B:PROD.0000012454.06094.29.

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.

Burušić, J., Babarović, T., & Velić, M. Š. (2016). School effectiveness: An overview of conceptual, methodological and empirical foundations. In *School effectiveness and educational management* (pp. 5–26). Springer.

Camanho, A. S., Varriale, L., Barbosa, F., & Sobral, T. (2021). Performance assessment of upper secondary schools in Italian regions using a circular pseudo-Malmquist index. *European Journal of Operational Research, 289*(3), 1188–1208.

Cazals, C., Florens, J.-P., & Simar, L. (2002). Nonparametric frontier estimation: A robust approach. *Journal of econometrics, 106*(1), 1–25.

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

Cherchye, L., Moesen, W., Rogge, N., & Puyenbroeck, T. V. (2007). An introduction to ‘benefit of the doubt’ composite indicators. *Social Indicators Research, 82*(1), 111–145. https://doi.org/10.1007/s11205-006-9029-7.

Clements, B. (2002). How efficient is education spending in Europe. *European Review of Economics and Finance, 1*(1), 3–26.

Council of the European Union (2009). Council conclusions of 12 may 2009 on a strategic framework for European cooperation in education and training (ET 2020). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52009XG0528(01).

Council of the European Union (2015). Joint report of the council and the commission on the implementation of the strategic framework for European cooperation in education and training (ET 2020) - New priorities for European cooper-

ation in education and training. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52015XG1215(02)&from=EN.
De Witte, K., & Hudrlikova, L. (2013). What about excellence in teaching? A benevolent ranking of universities. *Scientometrics, 96*(1), 337–364. https://doi.org/10.1007/s11192-013-0971-2.
De Witte, K., & López-Torres, L. (2017). Efficiency in education: A review of literature and a way forward. *Journal of the Operational Research Society, 68*(4), 339–363.
De Witte, K., & Rogge, N. (2010). To publish or not to publish? On the aggregation and drivers of research performance. *Scientometrics, 85*(3), 657–680. https://doi.org/10.1007/s11192-010-0286-5.
De Witte, K., & Schiltz, F. (2018). Measuring and explaining organizational effectiveness of school districts: Evidence from a robust and conditional benefit-of-the–doubt approach. *European Journal of Operational Research, 267*(3), 1172–1181.
European Commission (2015). *Education and training monitor 2015. Country analysis*. Luxembourg: Publication Office of the European Union.
European Commission (2017). *Education and training monitor 2017. Country analysis*. Luxembourg: Publication Office of the European Union.
European Commission (2018a). *Education and training monitor 2018. Country analysis*. Luxembourg: Publication Office of the European Union.
European Commission (2018b). *Education and training monitor 2018. Luxembourg*. Luxembourg: Publication Office of the European Union.
European Commission (2020a). *Education and training monitor 2020. Estonia*. Luxembourg: Publication Office of the European Union.
European Commission (2020b). *Education and training monitor 2020. Greece*. Luxembourg: Publication Office of the European Union.
European Commission (2020c). *Education and training monitor 2020. Ireland*. Luxembourg: Publication Office of the European Union.
European Commission (2020d). *Education and training monitor 2020. Latvia*. Luxembourg: Publication Office of the European Union.
European Commission (2020e). *Education and training monitor 2020. Luxembourg*. Luxembourg: Publication Office of the European Union.
European Commission (2020f). *Education and training monitor 2020. Malta*. Luxembourg: Publication Office of the European Union.
European Commission (2020g). *Education and training monitor 2020. Portugal*. Luxembourg: Publication Office of the European Union.
European Commission (2021). Education and training 2020. https://ec.europa.eu/eurostat/web/education-and-training/eu-benchmarks.
European CommissionEACEA and Eurydice (2015). *Structural indicators for monitoring education and training systems in Europe – 2015. Eurydice background report to the education and training monitor 2015. eurydice report*. Luxembourg: Publication Office of the European Union.
European CommissionEACEA and Eurydice (2016). *Structural indicators for monitoring education and training systems in Europe – 2016. Eurydice background report to the education and training monitor 2016. eurydice report*. Luxembourg: Publication Office of the European Union.
European CommissionEACEA and Eurydice (2019). *Key data on early childhood education and care in Europe – 2019 edition. Eurydice report*. Luxembourg: Publication Office of the European Union.
European CommissionEACEA and Eurydice (2020). *Structural indicators for monitoring education and training systems in Europe – 2020: Overview of major reforms since 2015. eurydice report*. Luxembourg: Publication Office of the European Union.
Färe, R., Grosskopf, S., Lindgren, B., & Roos, P. (1989). Productivity developments in Swedish hospitals: A Malmquist output index approach. In A. Charnes, W. Cooper, A. Lewin, & L. Seiford (Eds.), *Data envelopment analysis: Theory, methodology, and applications* (pp. 253–272). Academic Publishers. Reprinted by Springer in 1994.
Färe, R., Grosskopf, S., Lindgren, B., & Roos, P. (1992). Productivity change in Swedish pharmacies 1980–1989: A nonparametric Malmquist approach. *Journal of Productivity Analysis, 3*, 85–102.
Giambona, F., Vassallo, E., & Vassiliadis, E. (2011). Educational systems efficiency in European Union countries. *Studies in Educational Evaluation, 37*(2–3), 108–122.
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.
Giménez, V., Thieme, C., Prior, D., & Tortosa-Ausina, E. (2017). An international comparison of educational systems: A temporal analysis in presence of bad outputs. *Journal of Productivity Analysis, 47*(1), 83–101.
Giménez, V., Thieme, C., Prior, D., & Tortosa-Ausina, E. (2019). Comparing the performance of national educational systems: Inequality versus achievement? *Social Indicators Research, 141*(2), 581–609.
Hanushek, E. A., & Luque, J. A. (2003). Efficiency and equity in schools around the world. *Economics of Education Review, 22*(5), 481–502.
Horta, I. M., & Camanho, A. S. (2015). A nonparametric methodology for evaluating convergence in a multi-input multi-output setting. *European Journal of Operational Research, 246*(2), 554–561. https://doi.org/10.1016/j.ejor.2015.05.015.
Koopmans, T. (1951). An analysis of production as an efficient combination of activities. In T. Koopmans (Ed.), *Activity analysis of production and allocation* (pp. 33–97). The University of Chicago, New York: Wiley and Sons.
Le, M. H., Afsharian, M., & Ahn, H. (2021). Inverse frontier-based benchmarking for investigating the efficiency and achieving the targets in the vietnamese education system. *Omega, 103*, 102427.
López-Torres, L., & Prior, D. (2020). Long-term efficiency of public service provision in a context of budget restrictions. an application to the education sector. *Socio-Economic Planning Sciences*, 100946.
Lovell, C. K., Pastor, J. T., & Turner, J. A. (1995). Measuring macroeconomic performance in the OECD: A comparison of european and non-european countries. *European Journal of Operational Research, 87*(3), 507–518.
O'Donnell, C. J., Rao, D. S. P., & Battese, G. E. (2008). Metafrontier frameworks for the study of firm-level efficiencies and technology ratios. *Empirical Economics, 34*(2), 231–255. https://doi.org/10.1007/s00181-007-0119-4.
OECD (2020). *Productivity measurement and analysis* (1st ed.). Norway: OECD Publishing. http://www.oecd.org/education/policy-outlook/country-profile-Norway-2020.pdf.
Pastor, J. T., & Lovell, C. K. (2005). A global Malmquist productivity index. *Economics Letters, 88*(2), 266–271. https://doi.org/10.1016/j.econlet.2005.02.013.
Pereira, M. A., Camanho, A. S., Marques, R. C., & Figueira, J. R. (2021). The convergence of the world health organization member states regarding the united nations' sustainable development goal 'good health and well-being'. *Omega, 104*, 102495.
Rogge, N. (2019). EU countries' progress towards 'Europe 2020 strategy targets'. *Journal of Policy Modeling, 41*(2), 255–272. https://doi.org/10.1016/j.jpolmod.2019.03.003.
Silova, I., Millei, Z., Chachkhiani, K., Palandjian, G., & Vitrukh, M. (2021). Post-socialist transformations in comparative international education: Monuments, movements, and metamorphoses. In *The bloomsbury handbook of theory in comparative and international education*.
Stumbriene, D., Camanho, A. S., & Jakaitiene, A. (2019). The performance of education systems in the light of Europe 2020 strategy. *Annals of Operations Research*, 1–32.
Stumbriene, D., Zelvys, R., Zilinskas, J., Dukynaite, R., & Jakaitiene, A. (2022). Efficiency and effectiveness analysis based on educational inclusion and fairness of european countries. *Socio-Economic Planning Sciences*, 101293. https://doi.org/10.1016/j.seps.2022.101293.
Sutherland, D., Price, R., & Gonand, F. (2010). Improving public spending efficiency in primary and secondary education. *OECD Journal: Economic Studies, 2009*(1), 1–30.
Thieme, C., Giménez, V., & Prior, D. (2012). A comparative analysis of the efficiency of national education systems. *Asia Pacific Education Review, 13*(1), 1–15.
Thompson, R. G., Langemeier, L. N., Lee, C.-T., Lee, E., & Thrall, R. M. (1990). The role of multiplier bounds in efficiency analysis with application to Kansas farming. *Journal of Econometrics, 46*(1), 93–108. https://doi.org/10.1016/0304-4076(90)90049-Y.
Van Puyenbroeck, T. (2018). On the output orientation of the benefit-of-the-doubt–model. *Social Indicators Research, 139*(2), 415–431.
Wilson, P. W. (1995). Detecting influential observations in data envelopment analysis. *Journal of Productivity Analysis, 6*(1), 27–45.
Wüst, C., & Rogge, N. (2021). How is the european union progressing towards its europe 2020 targets? A benefit-of-the-doubt window analysis. *Empirica*, 1–34.
Zanella, A., Camanho, A. S., & Dias, T. G. (2015). Undesirable outputs and weighting schemes in composite indicators based on data envelopment analysis. *European Journal of Operational Research, 245*(2), 517–530. https://doi.org/10.1016/j.ejor.2015.03.036.
Želvys, R., Stumbriene, D., & Jakaitiene, A. (2018). Re-contextualization of effectiveness and efficiency in post-socialist education. *Bulgarian Comparative Education Society*.