Untitled

Bennour Mohamed Hsin

2024-02-04

Abstract

Education is the backbone of long-term economic prosperity. Thus, studying the educational landscapes of different countries for comparative reasons allows researchers to derive best practices that can be moulded into policies for the benefit of the collective well-being of communities in the regions. In this context, we conduct a detailed examination of school efficiency across 1190 schools at the NUTS2 regions of the Visegrad countries, namely Hungary, Poland, the Czech Republic, and the Slovak Republic, utilizing the rich dataset provided by PISA 2019. The dataset compiled for this research consists of an integrated mix of data, where student-level performance metrics and scores are aggregated to reflect school-level outcomes. This comprehensive dataset includes detailed information about the internal environment of each school, as well as the external factors surrounding it. Our analysis is threefold: first, we employ Data Envelopment Analysis (DEA) that identifies key factors contributing to school effectiveness and best practices in terms of student performancces; second, we employ a double-bootstrap procedure to estimate the contribution of various non-discretionary variables on school efficiency (related to the schools and to the region in question); and third, we conduct a Meta-Efficiency analysis across the four countries to unravel hidden patterns of inefficiencies. Our approach provides comprehensive analysis and insights that have actual practical policy implications, which can be tailored to the regional context of each of the countries in our study.

Introduction

Knowledge plays a pivotal role in shaping a nation's economic landscape, driving growth and prosperity [@hanushek2020_a] and [@hanushek2020_b]. It is considered a major factor for enhancing productivity and influencing a nation's overall economic situation [@jdadams1990]. The increasing of a nation's knowledge stock is closely related to the increasing of its human capital, achieved through investments in technology and training [@inglesi-lotz_2013] and [@hanushek2020_b]. Additionally [@hanushek2021] determined that the overall prevalence of patience and risk-taking practices across countries accounts for nearly two thirds of the variations that drives this investment in Human Capital, which is to say that there are specific dynamics that govern these mechanics of influence. This deliberate effort in investing in the Human Capital results in a multiplier effect for the economies in question, fostering innovation, encouraging healthy competition, and contributing to comprehensive growth and development [@inglesi-lotz_2014] and [@pmromer1989], this was also emphasized by [@cunha2006] who explicitly mentioned this multiplier effect of skill gains. The emphasis on knowledge (tacit knowledge, know-how, skills of different dimensions) has consistently been associated with a nation's education system and its quality, strongly linked to long-term economic growth development and overall prosperity, for instance [@hanushek2012] derived a causal effect of cognitive skills of a population on the long-term economic growth rates across countries.However, sustaining this correlation becomes a challenge as the socio-economic contexts and conditions change and evolve [@hanushek2016]. Understanding the mechanics that drives investment in Human Capital, and the causal effect of cognitive skills on economic growth are both a great starting point to address what constitutes these skills, and how they are acquired, by whom, and more imortantly how can agents of interest in this context optimize their return on investment and make sure that the resources enabled for this process are fully utilized, hence the focus on education, and educational institutions. The literature provided various scopes on how to approach educational landscapes from this perspective, two main scopes are efficiency (doing the right things) and effectiveness (doing things right). Essentially effectiveness focus on the practices (resources, operations, pedagogy, etc)

that yields a desired result in terms of output, whereas efficiency (the scope of this paper) focuses on the operational dimension of these practices and how to maximize a desired output given a status quo of inputs. [@dewitte2017] in a comprehensive review criticizes the dychotomy between the literature on efficiency and effectiveness in the context of an educational landscape, arguing that both scopes have distinct contributions that must be connected together for a rather comprehensive understanding of the issues challenges and adavantages of a given system. For the purposes of this paper we will approach the problem of assessing an educational landscape from the efficiency perspective, essentially given the finite nature of resources that are versed into a given system and because of the public nature of a majority of the education instituion which require additional attention to the policies implemented and their results and consequences on the public. This justifies the study of education systems as it allows to pinpoints the advantages and challenges that countries face in terms of educational excellence. following this line of thought, [@worthington2001] assumes that education institutions conduct a basic transformation of inputs into outputs, and this transformation can be proxied with a production function which can be estimated directly (parametrically) or indirectly (non-parametrically). Using Frontier estimation methods such as Data-envelopment-Analysis ([@charnes1978]) or Stochastic-Frontier-Analysis ([@aigner1977]) (or other similar methodologies) expanded the analysis tools that researchers used to approach educational institutions from the input/output perspective. Building on this, frontier methodologies are often utilized in a two-stage set up to further assess how efficient Decision Making Units (or DMU) react with other non-discretionary variables, this methodology has been observed in various studies including [@mayston_jesson1988], [@ray2004], and [@ruggiero2004] among others. For this paper we decided to adopt this approach where we estimate (non-parametrically) the production function of our school sample with the means of Data-envelopment-analysis (or DEA, yielding efficiency scores) in the first stage then we use a a double-bootstrap process (following [@sw2007]) to estimate a truncated regression that enables inference on some non-discretionary factors within the DMUs and around it. In this study we will use a sample of schools spanning across the Visegrad countries (V4:) extracted from the PISA survey data of 2022. This data allows us to specify a set of inputs, outputs, and non discretionary factors that we will use to derive best practices, and to derive inference in the second stage of our analysis. The remaining of the paper is as follows: section 2 will introduce the methodology and data used in the study, section 3 will provide the results, section 4 will present a discussion, and finally we conclude in section 5.

Data and Methodology

The literature is rich with studies regarding assessing educational landscapes of regions or cities (example: [@Butler_Monk_1985], [@Sengupta_Sfeir_1986], [@Barrow_1991], [@Ruggiero_1996a], [@Fukuyama_Weber_2002], [@Banker_et_al_2004], [@Davutyan_et_al_2010], and [@Johnson_Ruggiero_2014]), as well as countries (example: [@Gershberg_Schuermann_2001], [@Hanushek_Luque_2003], [@Afonso_Aubyn_2006], [@Kocher_et_al_2006], [@Gimenez_et_al_2007], [@Agasisti_2011b], [@Thieme_et_al_2012], and [@Aristovnik_2013]). There are also studies that tackles the assessment process from other angles, for example [@Cooper_Cohn_1997] (used data on classes in South Carolina) and [@De_Witte_Rogge_2011] (assessed teachers' performance across various college courses). Additionally other papers studied and assessed the student level performance, including [@Thanassoulis_1999], [@Colbert_et_al_2000], [@Thanassoulis_Portela_2002], [@Dolton_et_al_2003], [@Johnes_2006b], [@Deutsch_et_al_2013], [@Portela_et_al_2013], [@Thieme_et_al_2013], [@Crespo-Cebada_et_al_2014], and [@Podinovski_et_al_2014] to cite a few. Another scope of analysis remains, and that's the school level assessment which can be found in [@Bessent_Bessent_1980], [@Charnes_et_al_1981], [@Bessent_et_al_1982], [@Ray_1991], [@Thanassoulis_1996], [@Bradley_et_al_2001], [@Ouellette_Vierstraete_2005], [@Grosskopf_et_al_2009], [@Essid_et_al_2010], [@Agasisti_2011a], [@Johnes_et_al_2012], [@Burney_et_al_2013], and [@Brennan_et_al_2014] among others.

Moreover, in order to assess these educational institutions (at any level of analysis), different methodologies can be used and employed depending on the premis, the need, and the objective of the study. In the realm of frontier analysis, DEA is the most common methodology used in such papers, however, DEA is commonly combined with other methodologies as well such as Bootstrapping [@essid2010; @essid2013], or simply a regression procedure (be it ols, tobit or truncated) [@ruggiero1998; @afonso2006; @agasisti2007; @johnes2012] or even using the Malmquist index [@worthington2008; @ouellette2010; @essid2014] to cite a few. [@dewitte2017] provides one of the most comprehensive reviews of the literature in this regard.

In this paper we will adopt the school level analysis scope. The assumption that Educaitional institutions can be regarded as productive units that transforms inputs into outputs ([@worthington2001]), enables researchers to scrutinize data on school level with flexibility regarding the scope of the study or the angle of investigation. For this reason, efficiency analysis, especially DEA, has been quite popular in the literature as it captures the mechanics of this said transformation by estimating the frontier of efficient DMUs (in which the deviation from the efficiency frontier is assumed to be entirely due to inefficiency). However, these institutions, schools in particular, are not isolated entities. In fact, there's multiple factors that affect a school's performance spanning from the individual circumstances of the students, teachers, and staff of that school, to the environmental factors surrounding it, from infrastructure, location, and other competing schools.

In order to conduct the investigation on the influential environmental factors on the efficiency of the schools in the Visegrad countries, we will leverage the PISA survey data from 2022 in a two-stage analysis, where in the first stage we estimate the efficiency of our units in each country, then we run a double bootstrap procedure that estimates the influence of each environmental factor in a truncated regression. In this section we will delve into more details on the data we used, the approach we adopt, and the model specifications.

Data

The PISA survey data is commonly used by researcher as it contains comprehensive information on many levels for multiple countries within the OECD and outside it. The survey contains multiple response files, on Student level, School level, and Teacher level (among others). The popularity of the use of this PISA survey data is reflected by the amount of papers published containing 'PISA data' or 'PISA survey' as a keyword which amounts to 322 papers from the Web of Science database and 414 from the Scopus database, between 2003 and 2024. In our study we combine a set a variables on school level and student level to build a comprehensive and aggregated school level data set that contain enough information that can be leveraged by the methodology we use to investigate the effect of non-discretionary variables on school efficiency for the Visegrad countries (Hungary (HUN), Poland (POL), Czech Republic (CZE), Slovak Republic (SVK)). Our final clean dataset contains 1180 observations originally, and 1085 observation after removing the missing values. The decomposition of the final dataset by country is described in table 1.

CNT	count
CZE	399
HUN	206
POL	234
SVK	246

Table 2 provide a full description of the final

variable	min	max	mean	sd
SC001Q01TA	1.0000	5.0000	2.822120	1.1439997
SC011Q01TA	1.0000	3.0000	1.546544	0.8109602
SCHSIZE	9.0000	1910.0000	461.549309	261.1375318
SC002Q01TA	0.0000	1327.0000	232.341014	168.1029815
SC002Q02TA	0.0000	1222.0000	229.208295	157.4050259
STRATIO	1.0000	100.0000	11.852931	4.4305150
TOTAT	4.0000	126.5000	39.494009	21.1891458
student_behavior_issue_mean	1.0000	4.0000	1.853641	0.5476199
teacher_behavior_issue_mean	1.0000	4.0000	1.746544	0.5125901
resources_issue_mean	1.0000	4.0000	1.982949	0.6449143
staff_issues_mean	1.0000	3.7500	1.805607	0.6801135
total_homework_time	1.0000	6.0000	2.436100	0.6977714
total_class_periods	3.0000	53.0000	29.490621	6.1199137

mean_PV_science	237.7939	675.7673	488.678950	70.2160829
mean_PV_reading	235.2858	664.2186	476.599113	73.7625559
mean_PV_math	265.3644	661.3145	479.648252	69.0450507
mean_reading_attitude	1.0000	4.0000	2.632934	0.2510639
mean_science_attitude	1.0000	4.0000	2.615239	0.2982299
mean_math_attitude	1.0000	4.0000	2.506144	0.2993757

CNT	measure	SC001Q01TA	SC011Q01TA	SCHSIZE	SC002Q01TA	SC002Q02TA	
CZE	min	1.0000000	1.0000000	13.0000	0.0000	4.0000	
CZE	max	5.0000000	3.0000000	1910.0000	1258.0000	1063.0000	
CZE	mean	2.7318296	1.6165414	459.7820	232.2406	227.5414	
CZE	sd	1.1433726	0.8599562	244.1590	160.5287	142.1128	
HUN	min	1.0000000	1.0000000	9.0000	3.0000	5.0000	
HUN	max	5.0000000	3.0000000	1461.0000	968.0000	1222.0000	
HUN	mean	3.1699029	1.6747573	497.2282	246.1893	251.0388	
HUN	sd	1.2198628	0.8358105	272.1139	168.4678	171.5327	
POL	min	1.0000000	1.0000000	23.0000	10.0000	0.0000	
POL	max	5.0000000	3.0000000	1428.0000	1327.0000	771.0000	
POL	mean	3.2094017	1.3119658	508.7137	258.9274	249.7863	
POL	sd	0.9373951	0.6359289	297.7659	207.6116	185.1188	
SVK	min	1.0000000	1.0000000	42.0000	1.0000	0.0000	
SVK	max	4.0000000	3.0000000	1344.0000	718.0000	671.0000	
SVK	mean	2.3089431	1.5487805	389.6748	195.6179	194.0569	
SVK	sd	1.0311878	0.8154494	224.0781	127.4289	132.4911	

Inputs Choosing the inputs that will be used in our analysis is very important for few reasons. First, in the context of analyzing schools' efficiency, the inputs of schools are almost fixed, thus we have to use plausible inputs with enough variation. We decided to use as inputs the students to teacher ratio (STRATIO) which has been used in other studies such as [@Bessent_Bessent_1980; @Charnes_et_al_1981; @Ray_1991; @Johnes_1996; @Cooper_Cohn_1997; @Heshmati_Kumbhakar_1997; @Afonso_Aubyn_2006; @Johnes_Yu_2008; @Cherchye_et_al_2010; @Agasisti_2011b; @Agasisti_2013; @Johnes_2013; @Zoghbi_et_al_2013] among others and the total number of schooling hours (total_class_period) which has been used in similar studies such as [@rebai2020; @benyahia2018] to cite a few.

CNT	measure	STRATIO	total_class_periods
CZE	min	3.250000	4.000000
CZE	max	100.000000	41.125000
CZE	mean	13.041722	30.157164
CZE	sd	5.566654	4.664055
HUN	min	1.000000	3.000000
HUN	max	25.408700	53.000000
HUN	mean	10.892397	29.720361
HUN	sd	3.212498	7.138557
POL	min	1.000000	3.000000
POL	max	20.192300	42.840000
POL	mean	10.093012	30.768631
POL	sd	3.437797	7.056765
SVK	min	3.142900	3.000000
SVK	max	21.615400	44.111111
SVK	mean	12.403189	27.001467

SVK	sd	3.208878	5.641514

Outputs   For the output we will adopt variables depicting aggregate (average) students' test scores by school, these aggregations describe the average scores students got for each school in MATH, READING, and SCIENCE. This is in accordance with multiple studies in this context such as[@Bessent_Bessent_1980; @Charnes_et_al_1981; @Afonso_Aubyn_2006; @Cherchye_et_al_2010; @Essid_et_al_2010; @Essid_et_al_2013; @De_Witte_Kortelainen_2013; @De_Witte_et_al_2013; @rebai2020; @benyahia2018]

CNT	measure	mean_PV_science	mean_PV_reading	mean_PV_math
CZE	min	237.79390	323.64400	349.40678
CZE	max	675.76735	664.21856	661.31452
CZE	mean	503.25716	493.41212	492.42704
CZE	sd	69.88695	68.93268	66.56293
HUN	min	322.45810	241.53110	265.36440
HUN	max	659.59434	628.47634	636.92900
HUN	mean	472.74453	458.12559	456.75072
HUN	sd	76.14176	83.91720	77.68297
POL	min	358.25930	235.28580	305.96565
POL	max	638.34327	627.14851	632.53511
POL	mean	500.72602	489.69615	489.46124
POL	sd	63.00998	71.32658	62.90285
SVK	min	294.74706	278.50268	282.69305
SVK	max	645.01216	648.83560	642.72645
SVK	mean	466.91785	452.34078	468.76174
SVK	sd	63.86884	64.19883	64.53600

Non-discretionary variables   The final piece of our data is the non-discretionary variables which we will use to assess the contribution of environmental variables in the schools and around them on the their efficiency. These variables are typically used in a two-stage analysis such as ours and provide insights that enables inference in order to understand what can be done (according to the sample of the study) to imptove the school efficiency that maximizes the output of test scores. We incorporated in the second stage of our analysis variables regarding behavior issues (absence, delinquency, harassment, violence, etc) of students (student_behavior_issue_mean), teachers (teacher_behavior_issue_mean), and staff (staff_issues_mean), as well as issues of available resources in the premis including personal and material resources (resources_issue_mean), number of schools in the area (SC011Q01TA, 3 none, 2 one school, 1 two ore more schools), Community type where the school is located (SC001Q01TA, 1 urban, 0 rural), school size depicted by the number of enrolled students in the school (SCHSIZE), and the total number of all teachers at the school (TOTAT).

CNT	measure	student_behavior_issue_mean	teacher_behavior_issue_mean	resources_issue_mean	staff_issues
CZE	min	1.0000000	1.0000000	1.0000000	1.
CZE	max	3.6000000	3.4000000	4.0000000	3.
CZE	mean	1.8952381	1.8616541	1.9555138	1.
CZE	sd	0.4739997	0.5009622	0.5805643	0.
HUN	min	1.0000000	1.0000000	1.0000000	1.
HUN	max	3.8000000	4.0000000	3.7500000	3.
HUN	mean	1.6495146	1.6436893	2.1771845	1.
HUN	sd	0.5715830	0.5007643	0.6704850	0.
POL	min	1.0000000	1.0000000	1.0000000	1.
POL	max	3.6000000	3.4000000	3.2500000	3.
POL	mean	2.0264957	1.6854701	1.8675214	1.
POL	sd	0.5258785	0.5311361	0.6211341	0.

		1.0000000	1.0000000	1.0000000	1.
SVK	min	1.0000000	1.0000000	1.0000000	1.
SVK	max	4.0000000	3.2000000	4.0000000	3.
SVK	mean	1.7926829	1.7040650	1.9745935	1.
SVK	sd	0.5961332	0.4907967	0.7098300	0.

Methodology

The choice of the methodology is directly linked to the objective of this study. Since our aim is to investigate the environmental factors that influence the efficiency of schools in our sample we will first start by determining how we will estimate this efficiency, an then we will need to estimate the contributions of the other factors. Thus a two-stage analysis in plausible for this purpose. This two-stage analysis is based on an initial estimation of the efficiency using DEA, which will be regresses on other variable in the second stage.

To execute our approach, we start by scrutinizing DEA, the result of the pioneering work of [@charnes1978] which was later popularized by [@banker1984], the idea is basically a linear program that estimates a set of parameters for each DMU that allows to rank these DMUs in order of their best practices or technical efficiency additionally depicting a frontier of best practice determined by efficient DMUs. DEA programs can be input oriented (minimize inputs at the same level of output, yielding efficiency scores between $]0, 1]$ where a score of 1 indicate a fully efficient unit (FE)) or ouput oriented (maximize outputs at the same level of inputs, yielding efficiency scores between $[1, +\infty[$, where a score of 1 indicate a fully efficient unit (FE)) while accounting for returns to scale as a constraint in the program. As for the second stage, it's also straight forward, as many authors advocate for the use of OLS or tobit regression such as [@chilingerian2004], [@ray2004], and [@ruggiero2004].

However, an interesting debate has took place regarding the statistical plausibility of second stage regression. On the one hand, advocates of this approach argue that the simplicity and versatility of the approach allows for valuable and practical insight. OLS or Tobit regressions are usually used in this context allowing for inference, in fact [@hoff2007] provided a review of the use of tobit and OLS regression in a second stage, and advocated for the use of tobit. On the other hand opponents to the approach advocate for a more careful specifications, since a two-stage analysis doesn't stem for a data generating process (DGP), thus it's quite difficult to understand what is exactly being estimated in the second stage. One of the earliest critics of this approach was [@grosskopf1996] who specified that other than the lack of a DGP, there's a possibility that these non-discretionary variables of the second stage can be correlated with the output efficiency of the first stage. Both [@grosskopf1996] and [@ramalho2010] argued against the use of linear models in the second stage. Building on this criticism, [@simar2007] was the first paper that coherently justified the use of a two-stage DEA proposing a DGP and two bootstrap algorithm that enable inference in the second stage with a truncated regression. The basic idea behind [@simar2007] work, is that DEA efficiency scores are non-parametrically estimated, thus they are a non-observed estimate of an unkown efficiency frontier which means that these scores do not account for stochasticity (i.e: any deviation from the efficiency frontier is entirely due to inefficiency) which means that using non-parametric methods in the second stage is more appropriate, hence the bootstrap algorithms.

In this study, we adopt an output oriented model with the aim to maximize the average test scores' results of the schools in our sample data with variable returns to scales for the first stage, whereas for the second stage, we adopt the second algorithm (Double Bootstrap) from [@simar2007] to estimate the parameters of the second stage truncated regression, at 95% confidence interval, with the first bootstrap loop at L1=100, and the second one at L2=2000.

The DEA program is given by the following:

For a set of n Decision Making Units (DMUs), the inputs and outputs form matrices where each DMU is associated with m inputs $X = [x_1, x_2, ..., x_m]$ and s outputs $Y = [y_1, y_2, ..., y_s]$ we optimize the following:

$$\text{Maximize} \quad \theta$$

$$\text{Subject to:} \quad \sum_{j=1}^{n} \lambda_j y_{rj} \geq \theta y_{ro}, \quad \forall r = 1, \ldots, s,$$

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

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

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

Where:

- θ is the efficiency score of DMU_0.
- x_{ij} and y_{rj} are the inputs and outputs of the jth DMU, respectively.
- λ_j are the weights assigned to each DMU, indicating the contribution of each DMU to the construction of the virtual DMU against which DMU_0 is being benchmarked.

Results

| CNT	mean_efficiency	sd_efficiency	min_efficiency	max_efficiency
CZE	1.269435	0.1496696	1	1.773932
HUN	1.252629	0.1723598	1	1.788242
POL	1.182705	0.1159448	1	1.575211
SVK	1.226372	0.1414102	1	1.734475

Distribution of Efficiency Scores by Country

Including Fully efficient schools
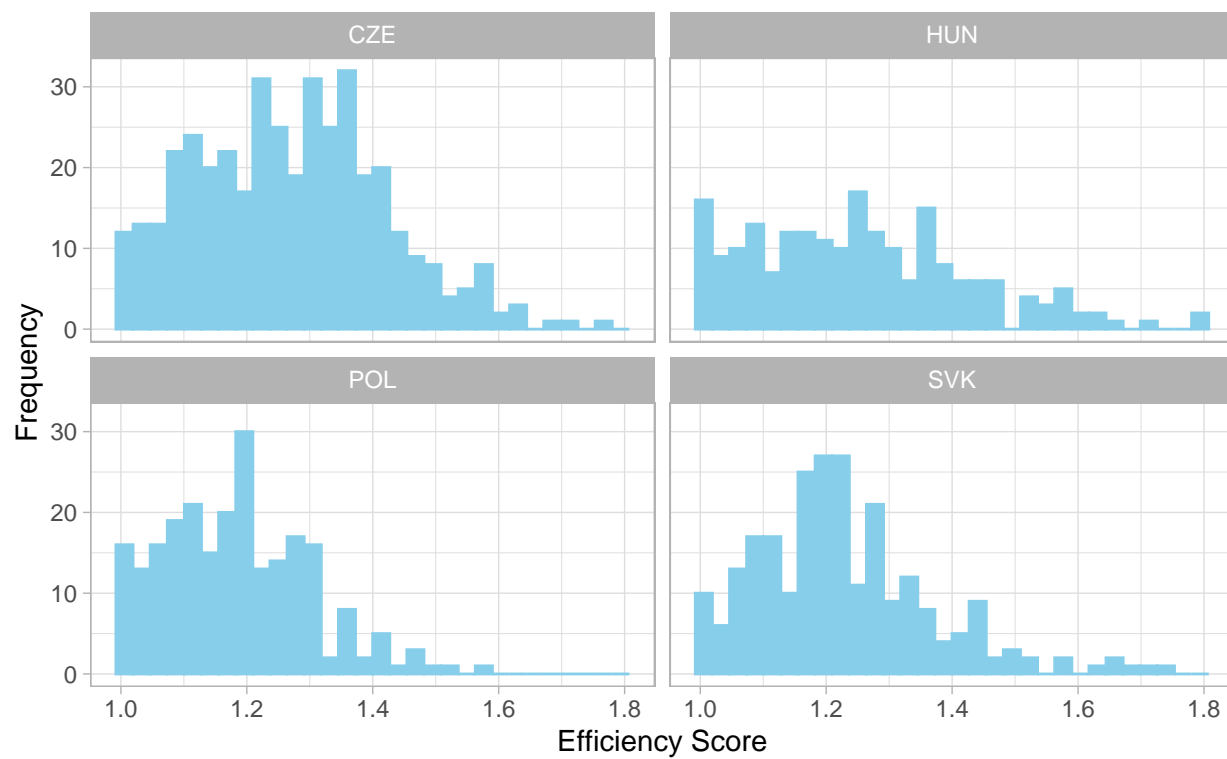
Distribution of Efficiency Scores by Country and Quartile
## For non Efficient Schools
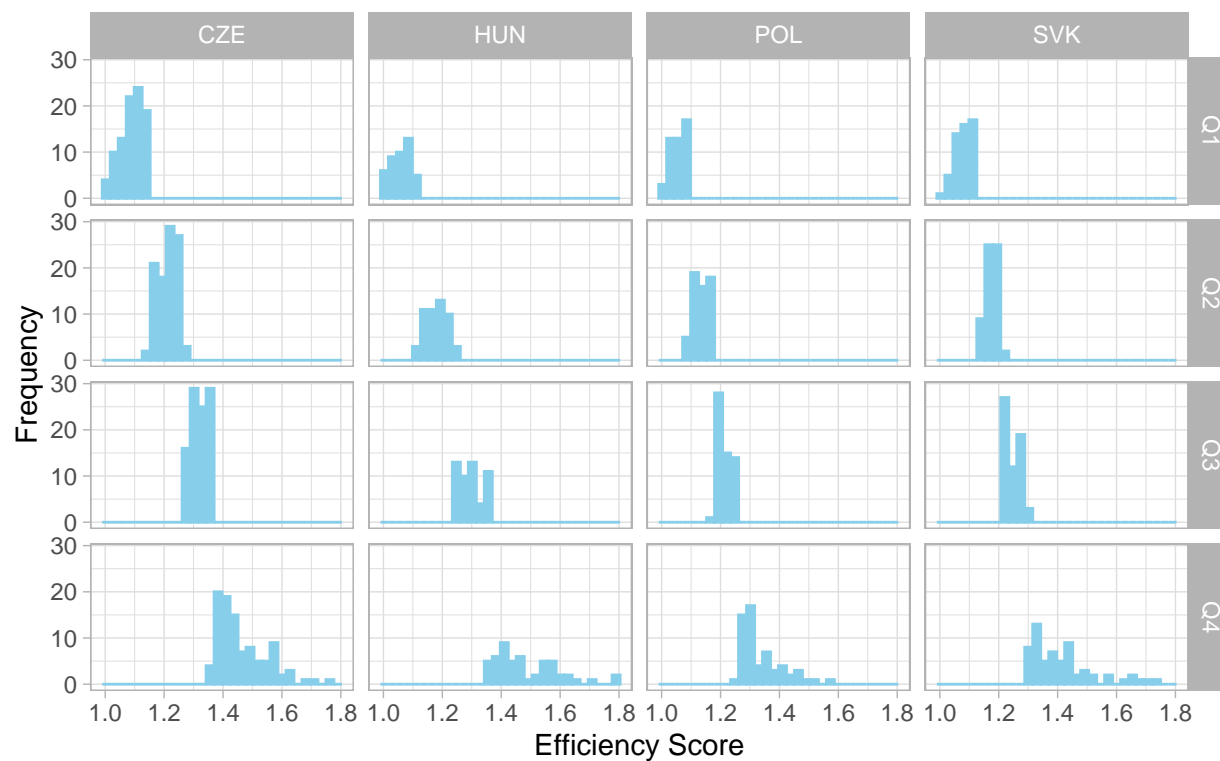
## Distribution of Efficiency Scores by Country and Quartile
For non Efficient Schools
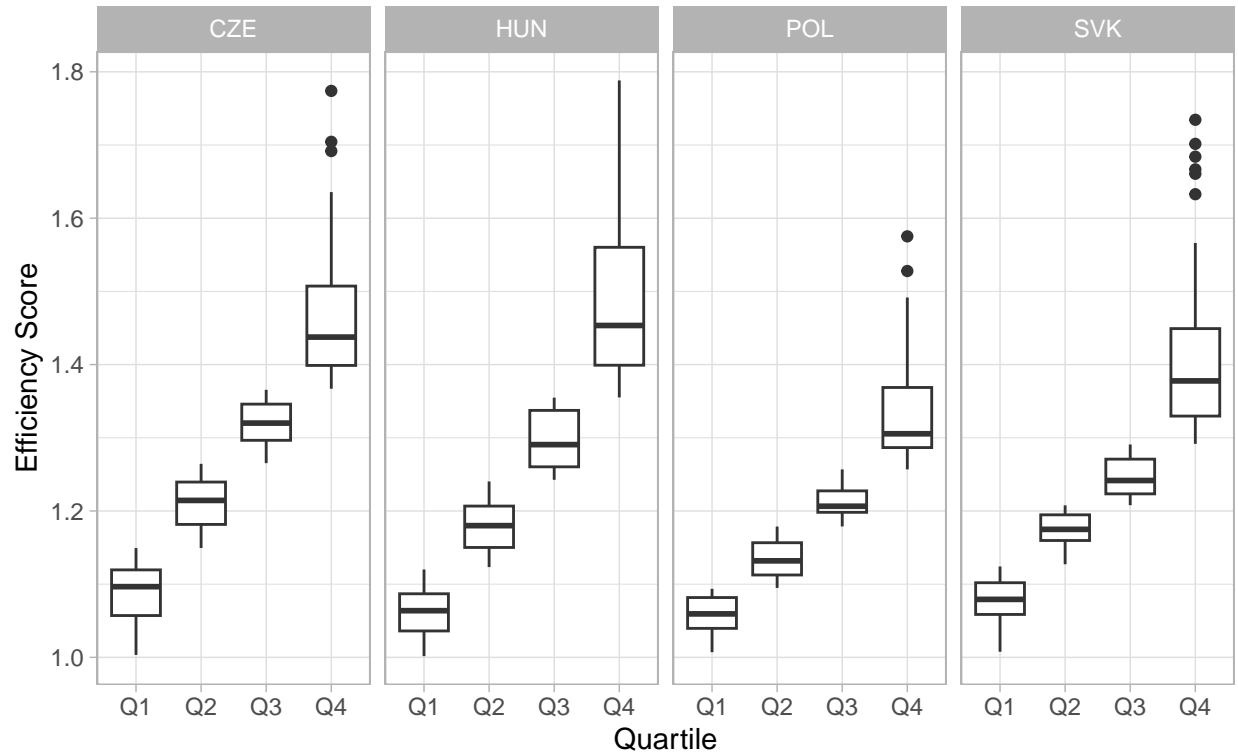
CNT	FE	Q1	Q2	Q3	Q4
CZE	8 (1.92%)	92 (22.06%)	100 (23.98%)	99 (23.74%)	100 (23.98%)
HUN	9 (3.54%)	43 (16.93%)	51 (20.08%)	51 (20.08%)	52 (20.47%)
POL	13 (5.46%)	46 (19.33%)	58 (24.37%)	58 (24.37%)	59 (24.79%)
SVK	9 (3.32%)	53 (19.56%)	61 (22.51%)	61 (22.51%)	62 (22.88%)

```
## Warning in geom_col(aes(x = n, y = CNT, .group = CNT, fill = quartiles_ineff),
## : Ignoring unknown aesthetics: .group
```

Percentage of Fully Efficient (FE) and Quartile distribution of Inefficient Schools by Country
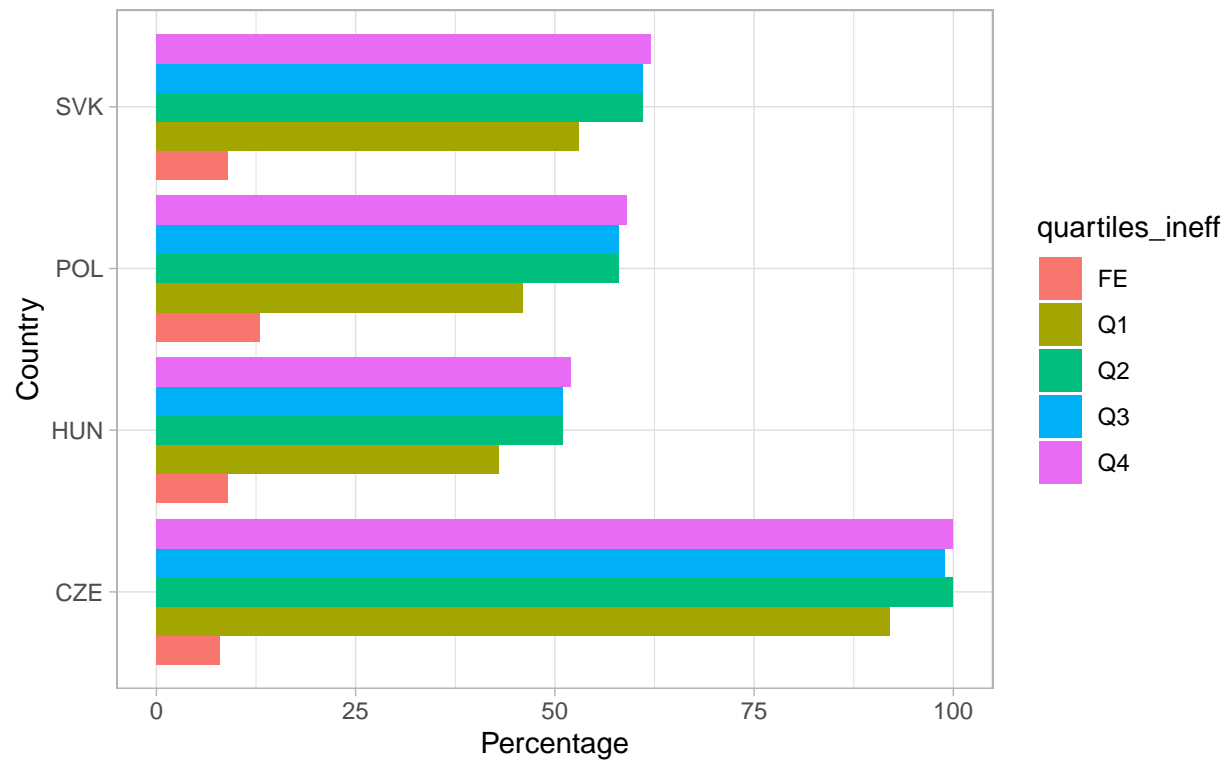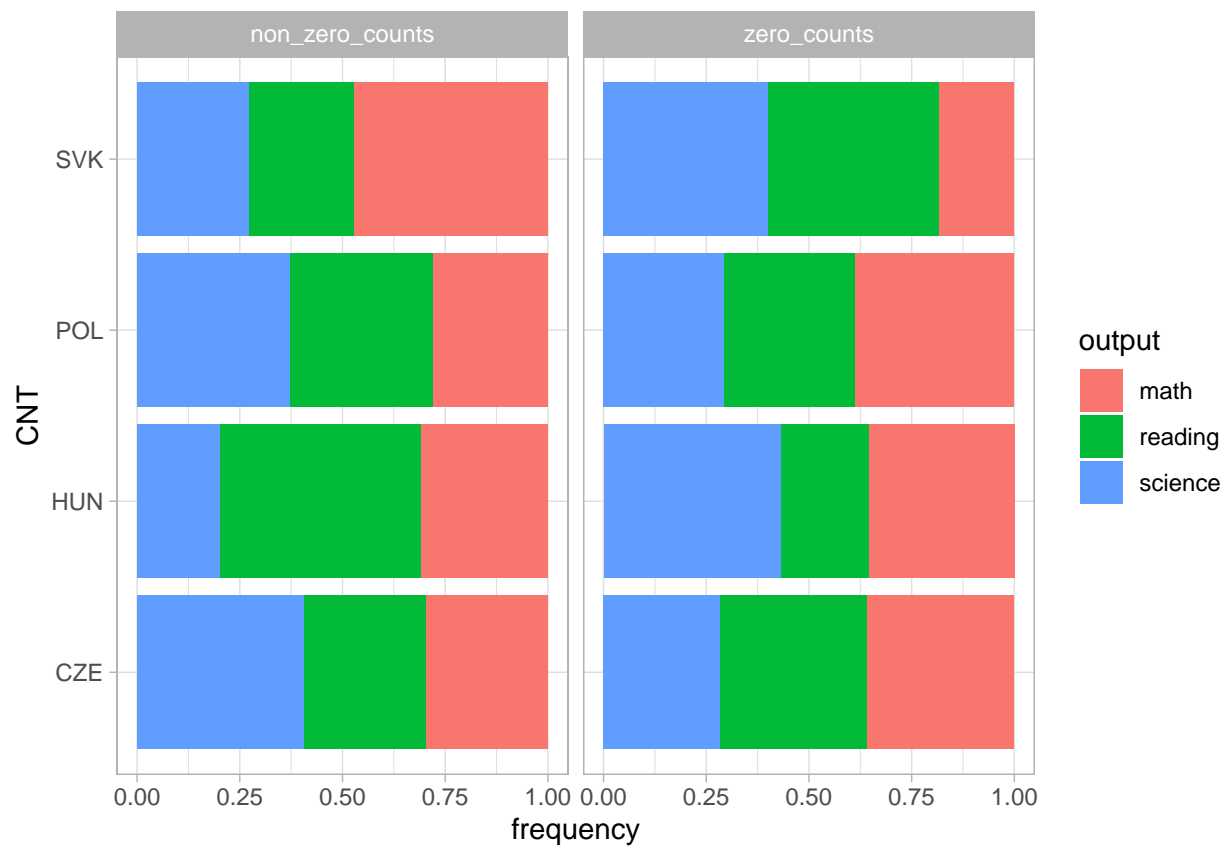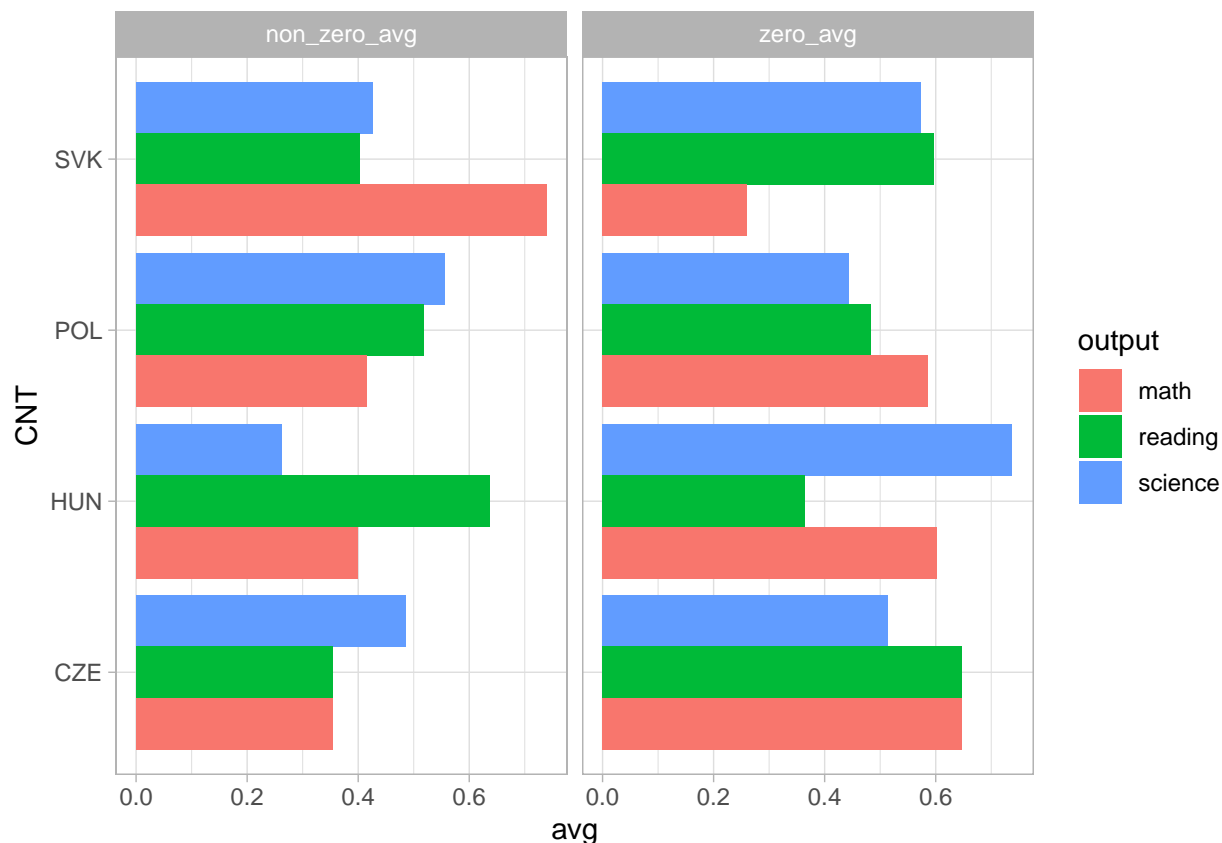
results of the efficiency, multipliers and the intensity weights

all the plots

Variable	CZE	HUN	POL	SVK
Intercept	1.626* (1.511, 1.742)	1.545* (1.376, 1.727)	1.568* (1.458, 1.673)	1.686* (1.541, 1.824)
Student Behavior Issue Mean	0.122* (0.092, 0.154)	0.227* (0.170, 0.283)	0.102* (0.067, 0.138)	0.097* (0.062, 0.134)
Teacher Behavior Issue Mean	-0.069* (-0.099, -0.039)	-0.053 (-0.110, 0.006)	-0.043 (-0.076, -0.010)	-0.044 (-0.088, -0.000)
Resources Issue Mean	0.032* (0.004, 0.059)	-0.046 (-0.099, 0.007)	-0.023 (-0.054, 0.009)	-0.001 (-0.034, 0.032)
Staff Issues Mean	-0.009 (-0.033, 0.016)	-0.010 (-0.060, 0.041)	-0.019 (-0.050, 0.014)	0.014 (-0.018, 0.046)
Total Class Periods	-0.015* (-0.018, -0.012)	-0.007* (-0.011, -0.004)	-0.010* (-0.013, -0.009)	-0.015* (-0.018, -0.011)
SCHSIZE	0.0002* (0.0001, 0.0003)	-0.0001 (-0.0003, 0.0001)	-0.0001* (-0.0002, -0.0000)	0.0002* (-0.0000, 0.0004)
SC001Q01TA	-0.016* (-0.030, -0.002)	-0.047* (-0.073, -0.022)	-0.030* (-0.047, -0.015)	-0.051* (-0.075, -0.029)
TOTAT	-0.003 (-0.004, -0.001)	-0.001 (-0.003, 0.002)	0.0004 (-0.001, 0.002)	-0.004* (-0.007, -0.001)

Discussion

limitations of the dea methodology and the specifications.

how these results can have policy implications

Conclusion

summarize.

References