Efficiency of Schools in the Visegrad Region, a StoNED Approach

BENNOUR Mohamed Hsin, University of Pécs

Abstract

The subject of school efficiency has attracted lots of literature attention, as school-level inefficiencies can have serious negative consequences on the development of human capital and the labor market. In this paper, we examine the efficiency of a sample of a combined 686 schools in the Visegrad countries using the 2018 PISA survey data. We employ an innovative non-parametric approach that combines both the benefits of Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA), namely the Stochastic Nonparametric Envelopment of Data (henceforth StoNED) developed in Kuosmanen and Kortelainen (2012). We construct four different models to control for return to scale, contextual (nondiscretionary) input variables, as well as the inclusion of undesirable variables, and accounting for multiple outputs. These models are applied country-wise and across all four countries to compare their efficiency frontiers and derive policy recommendations.

Introduction

In some empirical applications, researchers often seek to analyze the performance of certain Decision-Making Units (hereafter as DMU, eg: a school). These DMUs are often characterized by a production process that takes a certain level of resources (or inputs) and transforms them to a certain level of goods (or outputs). In this context researchers model productivity and efficiency as metrics of performance, by analyzing the production process which yields a production frontier. The frontier describes the feasible level of outputs for each level of inputs, this frontier reflects the state of technology for a given sector or industry.

Educational institutions can be considered (from an economic point of view) as production units (DMUs), and can be evaluated in terms of efficiency. Efficiency can be analyzed from various perspectives. School-level inefficiencies and mismanagement can have serious repercussions for human resource development and labor market outcomes as a recent study (Bhutoria and Aljabri 2022) highlights.

There are many papers that are measuring the efficiency within the sector of education, see for example the excellent review of Witte and López-Torres (2017). To the best of our knowledge, no paper among these concentrates on the Visegrad (V4) countries, although the countries are historically linked and followed similar development paths and it seems useful to compare their education policies and performance Pelle and Kuruczleki (2016). In this paper we focus on answering the question of what drives the efficiency of schools in the Visegrad countries in terms of standardized test performance? The efficiency analysis will be based on PISA 2018 data and the novel StoNED framework.

Methodology

The early literature provided two major directions in performance analysis. The first is the "Stochastic Frontier Analysis" (SFA). The second direction is the non-parametric approach, leading to the development of "Data Envelopment Analysis" (DEA) and "Full Disposal Hull" (FDH). DEA and FDH gained a large theoretical and empirical literature grounds since they were introduced (DEA and FDH are considered the same technique, they are usually both referred to as DEA).

Both approaches have clear advantages and disadvantages. The main shortcoming of DEA is that all deviations from the frontier are treated as inefficiency. The strength of SFA is its probabilistic nature. It allows for decomposing deviations into an inefficiency term and noise term. However development on convex nonparametric least squares (CNLS) have led to the full integration of DEA and SFA into a unified framework of productivity analysis, which is refered to as Stochastic Nonparametric Envelopment of Data (StoNED). Johnson and Kuosmanen (2015)

Formulation of a StoNED model requires identifying input, output and environmental variables, and solving an optimization algorithm, depending on various possible assumptions. In this paper we will use four variations of the StoNED Framework:

- Model 1 : CNLS estimation with contextual variables and constant returns to scale ($\beta_i = 1$).
  - one input, multiple outputs
- Model 2 : CNLS estimation with contextual variables and variable returns to scale ($\beta_i \geq 0$)
  - one input, multiple outputs

    the following Program(with VRS ($\beta_i \geq 0$)) represent the two models with the only difference being in the last constraint.

$$
\begin{aligned}
&\min_{\alpha,\boldsymbol{\beta},\varepsilon} \sum_{i=1}^{n} \varepsilon_i^2 \\
\boldsymbol{s.t.}\quad & y_i = \alpha_i + \boldsymbol{\beta}_i^{'}\boldsymbol{x}_i + \varepsilon_i && \forall \boldsymbol{i} \\
& \alpha_i + \boldsymbol{\beta}_i^{'}\boldsymbol{x}_i \le \alpha_j + \boldsymbol{\beta}_j^{'}\boldsymbol{x}_i && \forall \boldsymbol{i}, \boldsymbol{j} \\
& \boldsymbol{\beta}_i \ge \mathbf{0} && \forall \boldsymbol{i}
\end{aligned}
$$

- Model 3 : CNLS estimation with Directional Distance Function
  - multiple inputs and outputs
  - undesirable inputs

$$
\begin{aligned}
&\min_{\alpha,\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\delta},\varepsilon} \sum_{i=1}^{n} \varepsilon_i^2 \\
s.t.\quad & \boldsymbol{\gamma}_i^{'}\boldsymbol{y}_i = \alpha_i + \boldsymbol{\beta}_i^{'}\boldsymbol{x}_i + \boldsymbol{\delta}_i^{'}\boldsymbol{b}_i - \varepsilon_i && \forall i \\
& \alpha_i + \boldsymbol{\beta}_i^{'}\boldsymbol{x}_i + \delta_i^{'}\boldsymbol{b}_i - \boldsymbol{\gamma}_i^{'}\boldsymbol{y}_i \le \alpha_j + \boldsymbol{\beta}_j^{'}\boldsymbol{x}_i + \delta_j^{'}\boldsymbol{b}_i - \boldsymbol{\gamma}_j^{'}\boldsymbol{y}_i && \forall i, j \\
& \boldsymbol{\gamma}_i^{'}g^y + \boldsymbol{\beta}_i^{'}g^x + \boldsymbol{\delta}_i^{'}g^b = 1 && \forall i \\
& \boldsymbol{\beta}_i \ge \mathbf{0}, \boldsymbol{\delta}_i \ge \mathbf{0}, \boldsymbol{\gamma}_i \ge \mathbf{0} && \forall i
\end{aligned}
$$

  -

- Model 4 : CNLS estimation with Directional Distance Function
  - multiple inputs and outputs

$$
\begin{aligned}
&\min_{\alpha,\boldsymbol{\beta},\boldsymbol{\gamma},\varepsilon} \sum_{i=1}^{n} \varepsilon_i^2 \\
\boldsymbol{s.t.}\quad & \boldsymbol{\gamma}_i^{'}\boldsymbol{y}_i = \alpha_i + \beta_i^{'}\boldsymbol{x}_i - \varepsilon_i && \forall \boldsymbol{i} \\
& \alpha_i + \boldsymbol{\beta}_i^{'}\boldsymbol{x}_i - \boldsymbol{\gamma}_i^{'}\boldsymbol{y}_i \le \alpha_j + \boldsymbol{\beta}_j^{'}\boldsymbol{x}_i - \boldsymbol{\gamma}_j^{'}\boldsymbol{y}_i && \forall \boldsymbol{i}, \boldsymbol{j} \\
& \boldsymbol{\gamma}_i^{'}g^y + \boldsymbol{\beta}_i^{'}g^x = 1 && \forall \boldsymbol{i} \\
& \boldsymbol{\beta}_i \ge \mathbf{0}, \boldsymbol{\gamma}_i \ge \mathbf{0} && \forall \boldsymbol{i}
\end{aligned}
$$

  -

Table 1: Number of observations per country

country	n
CZE	333
HUN	238
POL	240
SVK	376

Moreover the 3rd and 4th models are basically the same with the only difference regarding the specification of the Directional Vector $g = (x, y, gx, gy)$, in model 3 we construct the vector to simultaneously increase outputs and decrease undesirable outputs $g_{m3} = (x, y, b, g_x = 0, g_y = 1, g_b = -1)$ while for model 4 we simply ignore the undesirable output specification as in $g_{m4} = (x, y, g_x = 0, g_y = 1)$

Kuosmanen and Johnson (2010) provide GAMS and MATLAB code originally, there exists a Python package Dai et al. (2021) for implementing different models which is what we used in this paper.

In the next chapter we use nonparametric envelopment techniques to estimate a production (education) frontier for each Visegrad country, and for the region as a whole across the four countries.

Data

Data is extracted from the 2018 PISA survey, it contains information on schools (and their environments), teachers, and students (and parental engagement). Data analysis and exploration, and visualisation was done in R, while the frontiers were estimated in Python. In this section we summarise and interpret our data mainly using graphical methods. We show the relationship of the variables we use later in the StoNED models.

Summary statistics of the data

Our sample is constituted of 623 schools from the V4 countries decomposed per the table below. Data is gathered on an individual level, so the first task was to aggregate that to a school level. This was done in two different ways: taking means and medians of the individual level results for each of the schools. Medians are more robust statistics of the central tendency, and not influenced so much by outliers. The models below were estimated using both summaries, but there were no big differences, so we just highlight results from the models using the mean.

The variables and their names we initially decided to use based on the existing literature are the following:

- Students:
  - total number of students: `total_sudents`
  - gender characteristics: `students_male (number of) & students_female (number of)`
  - parents: `ap_parental_eng` (proportion of parents who engage with the school staff regarding their children education)
  - `avg_math` ,`avg_read` , and `avg_science` are the average grades for each respective subject(math, reading and science) for students in each school of the sample
- Schools:
  - School characteristcs: `ratio_f2m (ratio of female to male students), ratio_pt2ft (part time to full time tachers), ratio_t2s (students per teacher), dropout_rate, funding (percentage of government funding)`
- Teachers:
  - number of full and part time teachers: `teachers_ft`, `teachers_pt`
  - `total_teachers` : Total number of teachers
- Context:
  - Location (`bol_locataion`): 1 - School is in Urban area, 0 - School is in Rural area
  - Extra activities (`bol_extra_acts`): 1 - Yes there are official extra activities in the curriculum, 0 - No
  - Competition (`bol_competition`): 1 - There's at least another school in the area, 0 - There's no other schools in the area
  - Career guidance (`bol_career_guidance`): 1 - There's an official career guidance curriculum in the school, 0 - There's not

Mean table for each variable per country is depicted in table 2.

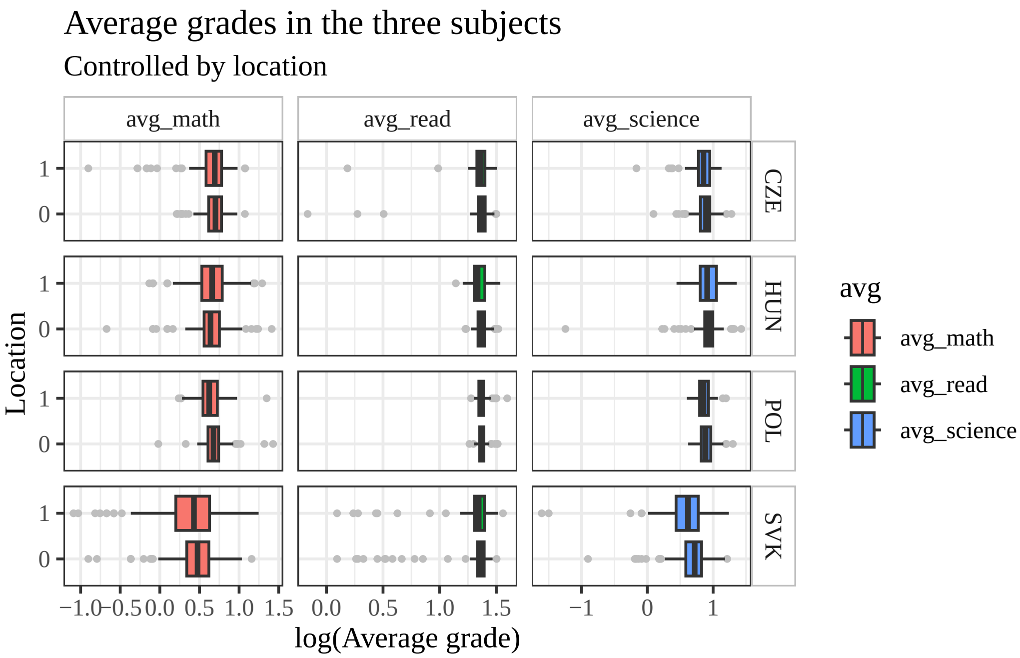
Based on the summary values we can see, that Poland has the highest proportion of urban schools in the sample (47%) whereas Hungary the lowest (only 21%). Official extra activities are so widely available that almost all institutions responded with yes, we can conclude that probably there is not enough variance in the data to find any effect in the models. Proportion of competing schools are similar across countries, however career guidance proportions are varying between 38% (Poland) and 72% (Czech Republic). Dropout rates are relatively low, but highest in the Czech Republic.

Table 2: Means of variables for each of the countries

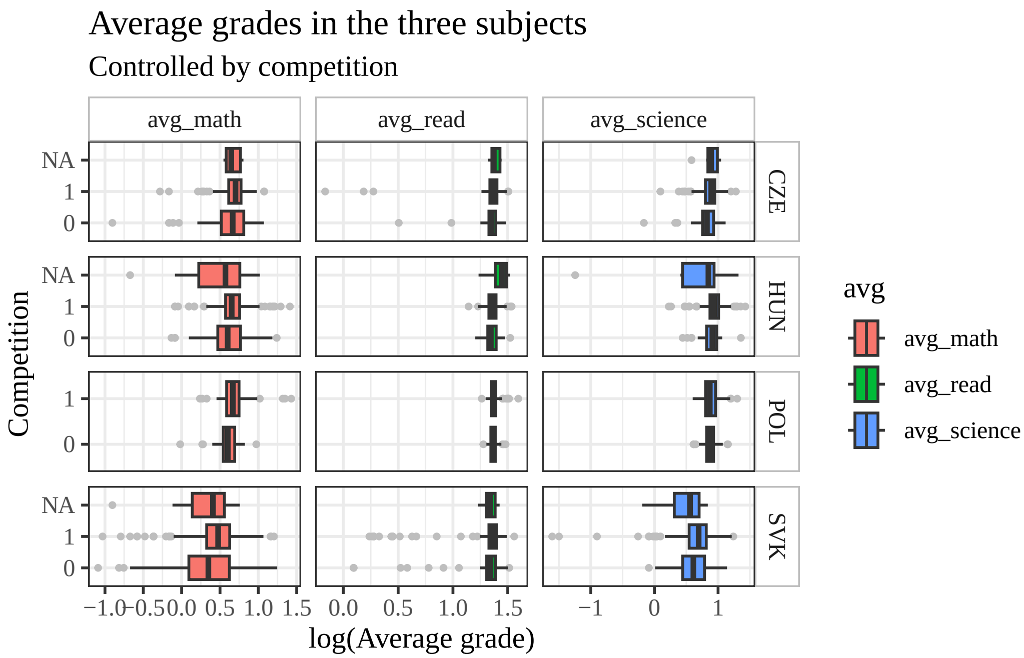
name	CZE	HUN	POL	SVK
bol_location	0.396	0.290	0.492	0.447
bol_extra_acts	0.994	0.996	1.000	1.000
bol_competition	0.830	0.809	0.729	0.821
bol_career_guidance	0.666	0.580	0.376	0.580
funding	0.954	0.959	0.968	0.969
students_male	5.188	5.245	4.834	4.950
students_female	5.111	5.094	4.812	4.877
teachers_ft	3.309	3.611	3.422	3.127
teachers_pt	1.685	1.264	1.952	1.355
total_students	5.932	5.949	5.536	5.696
total_teachers	3.551	3.743	3.703	3.332
ratio_f2m	-0.077	-0.150	-0.022	-0.073
ratio_ft2pt	1.642	2.396	1.480	1.788
ratio_s2t	2.396	2.216	1.833	2.399
ap_parental_eng	-1.463	-1.733	-1.073	-1.290
dropout_rate	-3.161	-3.546	-4.072	-3.604
avg_math	1.992	1.965	1.963	1.589
avg_read	3.909	3.917	3.955	3.794
avg_science	2.394	2.537	2.427	1.950

Visualisations

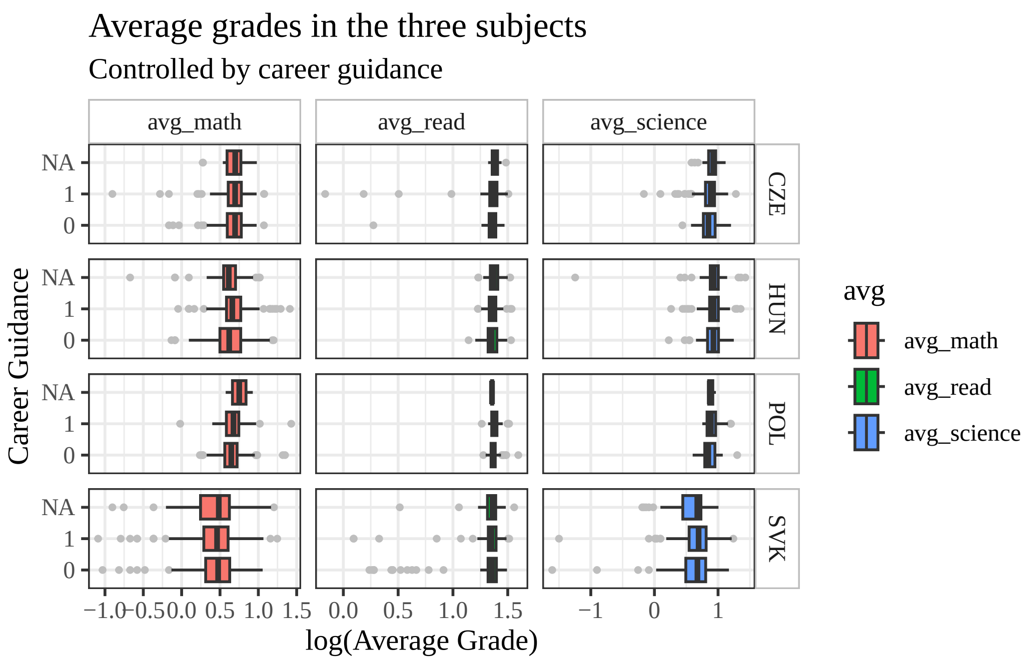
In this subsection we discover the patterns in the data using visual techniques. The following chart shows the average results in the three tests for each country, grouped by location. We can see some extreme outliers in almost all cases. Urban schools seem to slightly outperform rural schools in math scores in Hungary, but the opposite seems to be the situation in Slovak Republic. Interestingly science score on average are higher in rural located schools in general. Heterogeneity in terms of interquartile range is usually larger for urban schools, while rural school averages are more similar to each other. This seems to be the case for all V4 countries and all subjects. (Note that results are on a log scale!)



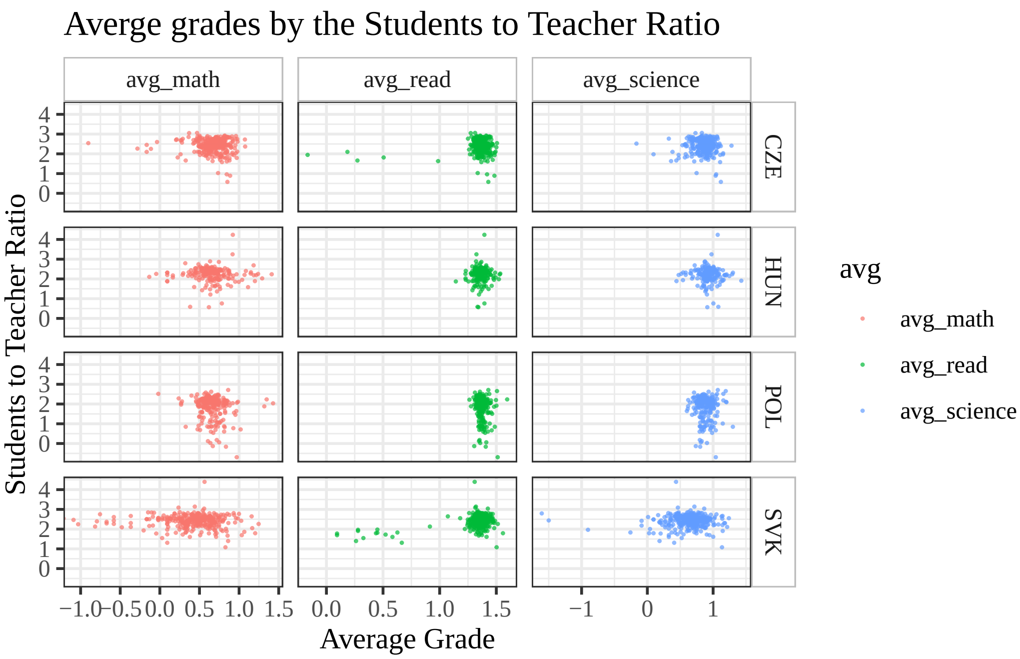
Next we checked for differences in terms of competition and performance. It seems to be evident that if institutions stated that there is competition, average results are better or similar. This conclusion holds through all countries and subjects as can be seen on the following graph.



Career guidance might improve grades on its own as well, but it also shows how important it is for the schools to develop their students' abilities according to their needs. Our data does not show a systematic relationship between performance (grades) and career guidance availability. Guidance might help and has an influence on later success of students, not the actual observable scores achieved.



Student to teachers ratios are one way to measure the quality of teaching. In general we observe low ratios in Polish schools, compared to the other three countries. In general, lower ratios indicate a somewhat better average result, however, the strength of the relationship seems to be weak.



Results and discussion

In this section we will elaborate on the different results of the models, and provide a brief discussion. However, it is important to note that not all StoNED models results are relevant, in fact in our case we focus mainly on assessing the inefficiencies and the contextual variables, other results such as the inputs or the outputs coefficients from the CNLS QP problem are not of interest in the case of the production frontier as its not clear on how to interpret them, but can be so in the case of cost frontier where they can be interpreted as elasticities. Also it's important to note that our approach for the the Visegrad sample data does not assume any differences between countries and thus no specific changes to the models were conducted. Sipilainen, Kuosmanen, and Kumbhakar (2008) provided a through explanation of the meta frontier approach with StoNED.

In the tables below country names are abbreviated to the ISO 3166method. As for the models, the model columns will contain the word `res` (for result) followed by the number of the model as per the methodology section.

Table 3: Summary statistics of firm specific inefficiencies by model and country

country	model	value_mean	value_sd
hun	res1	0.7438	0.0883
hun	res2	0.7818	0.0915
hun	res3	0.7262	0.0834
hun	res4	0.7189	0.0907
pol	res1	0.7365	0.1240
pol	res2	0.7870	0.0985
pol	res3	0.7459	0.0950
pol	res4	0.7416	0.1012
svk	res1	0.7099	0.1366
svk	res2	0.7769	0.2001
svk	res3	0.6707	0.1359
svk	res4	0.6650	0.1413
cze	res1	0.7424	0.1130
cze	res2	0.8206	0.0692
cze	res3	0.7467	0.0864
cze	res4	0.7449	0.0892

Inefficiencies

For the inefficiencies we can see in Table 3 the average school-specific deviation from the frontier per country and model. None of the schools are 100% efficient, which raises the question on how to optimize the quantifiable and observable processes to mitigate such issue. We can notice that for models 3 and 4 the average inefficiencies are almost similar which suggests that the incorporation of the Dropout rates as undesirable output(in model) might not have an effect on the inefficiencies especially as this variable is presented with low variance within, and across countries. However this still might be an interesting approach to investigate in the case of less developed countries such us the study in Yahia, Essid, and Rebai (2018).

We can also notice that model 2 (with variable returns to scale) always yields higher estimates of inefficiencies across countries than model 1 (with constant returns to scale)

The results associated with "meta" across country analysis in Table 4 also suggests that there isn't any clear difference between model 4 and 3 thus using the dropout rate as an undesirable output. Moreover, models 2 and 1 overestimate inefficiency with the presence of contextual variables.

Contextual variables

Table 4: Summary statistics of firm specific meta inefficiencies by model

model	mean	sd
res1_all	0.6750	0.1377
res2_all	0.7985	0.0706
res3_all	0.6139	0.1024
res4_all	0.6136	0.1028

Table 5: Contextual variables coefficients from Model 1 and 2 by country

country	model	z1	z2
svk	res1	3.2937	1.4066
svk	res2	-1.7916	-1.6127
hun	res1	7.3005	8.6229
hun	res2	2.6281	3.9037
cze	res1	11.5759	3.2709
cze	res2	2.9401	0.3056
pol	res1	5.7646	4.8820
pol	res2	1.3031	0.6774

The contextual variables employed in model 1 and 2 are the Booleans `Z1` : competition (to control the existence of other competing institutions in the same are) and `Z2` : location (to control for the effect of Urban vs Rural location).

In Table 5 we can see that model 1 in all cases overestimate the coefficients. This is also true in the case of the cross country models. In fact we can say that Competition can increase the Averaged grades by 9.7 points (model 1) or 1.04 points (model 2), additionaly Urban schools are expected to have an increase of 5.9 points (model1) or a decrease of 0.16 points (model 2).

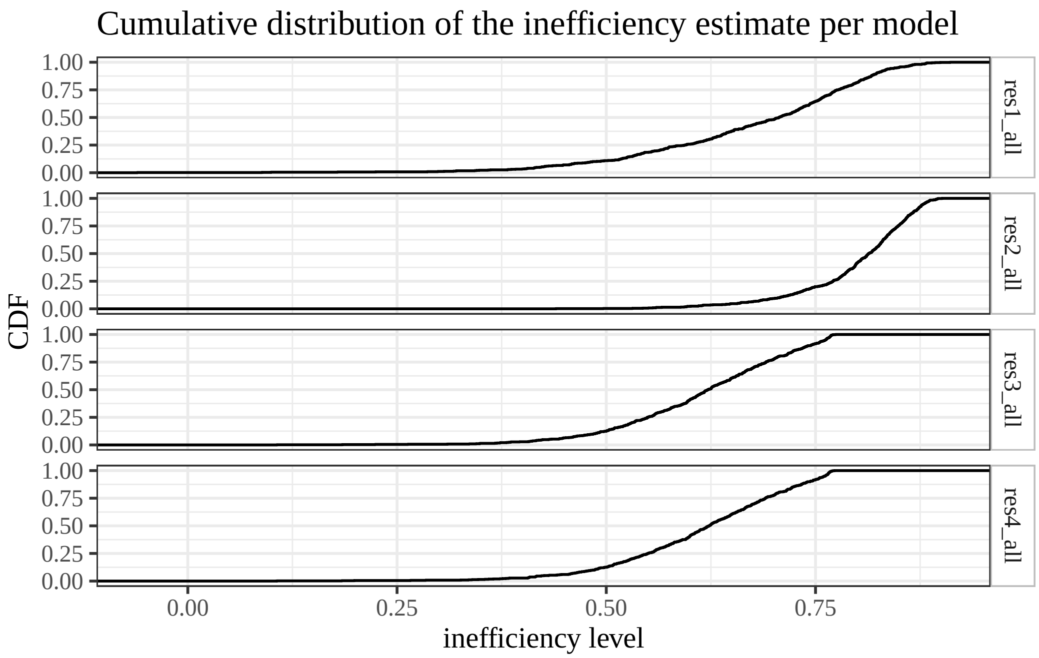
Of course these results seems inconsistent and might be the lead for further investigation especially for testing the returns to scale, constructing confidence intervals and significance levels of the variables.

Table 6: Contextual variables meta coefficients from Model 1 and 2 for all countries

model	z1	z2
res1_all	9.6508	5.8956
res2_all	1.0369	-0.1624

Cumulative meta-inefficiencies

following the spirit of Kuosmanen (2012) we constructed the cumulative inefficiency density plot in Figure, where we can see that most schools are highly inefficient despite the different set ups employed. Model 2 with VRS seems to overestimate efficiency.



Despite the use of various set ups and models, most Visegrad schools seem to be highly inefficient. The data we used summarises most the variance in this context as per the literature including Grosskopf, Hayes, and Taylor (2014) and Johnes, Portela, and Thanassoulis (2017). However more effort in capturing the anomalies in the data and determining more methodologically justified model specifications is required.

References

Bhutoria, Aditi, and Nayyaf Aljabri. 2022. "Managerial Practices and School Efficiency: A Data Envelopment Analysis Across OECD and MENA Countries Using TIMSS 2019 Data." *Large-Scale Assessments in Education* 10 (1): 24. https://doi.org/10.1186/s40536-022-00147-3.

Dai, Sheng, Yu-Hsueh Fang, Chia-Yen Lee, and Timo Kuosmanen. 2021. "pyStoNED: A Python Package for Convex Regression and Frontier Estimation." *arXiv Preprint arXiv:2109.12962*.

Grosskopf, Shawna, Kathy Hayes, and Lori L Taylor. 2014. "Applied Efficiency Analysis in Education." *Economics and Business Letters* 3 (1): 19–26.

Johnes, Jill, Maria Portela, and Emmanuel Thanassoulis. 2017. "Efficiency in Education." *Journal of the Operational Research Society*. Taylor & Francis.

Johnson, Andrew L., and Timo Kuosmanen. 2015. "An Introduction to CNLS and StoNED Methods for Efficiency Analysis: Economic Insights and Computational Aspects." In *Benchmarking for Performance Evaluation*, edited by Subhash C. Ray, Subal C. Kumbhakar, and Pami Dua, 117–86. Springer Books. Springer. https://doi.org/10.1007/978-81-322-2253-8.

Kuosmanen, Timo. 2012. "Stochastic Semi-Nonparametric Frontier Estimation of Electricity Distribution Networks: Application of the StoNED Method in the Finnish Regulatory Model." *Energy Economics* 34 (6): 2189–99.

Kuosmanen, Timo, and Andrew L Johnson. 2010. "Data Envelopment Analysis as Nonparametric Least-Squares Regression." *Operations Research* 58 (1): 149–60.

Kuosmanen, Timo, and Mika Kortelainen. 2012. "Stochastic Non-Smooth Envelopment of Data: Semi-Parametric Frontier Estimation Subject to Shape Constraints." *Journal of Productivity Analysis* 38: 11–28.

Pelle, Anita, and Éva Kuruczleki. 2016. "Education Policies and Performance of the Visegrad Countries in Light of Their OECD Membership: A Comparative Study." *Rocznik Instytutu Europy Środkowo-Wschodniej* 14 (4 Re-thinking the OECD's Role in Global Governance: Members, Policies, Influence): 173–207.

Sipilainen, Timo, Timo Kuosmanen, and Subal C Kumbhakar. 2008. "Measuring Productivity Differentials–an Application to Milk Production in Nordic Countries."

Witte, Kristof De, and Laura López-Torres. 2017. "Efficiency in Education: A Review of Literature and a Way Forward." *Journal of the Operational Research Society* 68 (4): 339–63. https://doi.org/10.1057/jors.2015.92.

Yahia, Fatma Ben, Hédi Essid, and Sonia Rebai. 2018. "Do Dropout and Environmental Factors Matter? A Directional Distance Function Assessment of Tunisian Education Efficiency." *International Journal of Educational Development* 60: 120–27.