Using PISA data for estimating the efficiency of secondary schools in an international perspective: preliminary results

Tommaso Agasisti* & Pablo Zoido

* **Politecnico di Milano School of Management**, Department of Management, Economics and Industrial Engineering.

OECD T.J. Alexander Fellow

e. tommaso.agasisti@polimi.it

Paper prepared for the Workshop "Efficiency in Education"
@ The Work Foundation, London, September 19th- 20th, 2014

Agenda

- Motivation and research questions
- Empirical model
  - Critical assumptions
- Data, descriptive statistics
- Results
  - Efficiency scores
  - Factors associated with efficiency scores
  - Robustness checks

Disclaimer

The author(s) received funding from the OECD Thomas J. Alexander fellowship program for carrying out this work. The work should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed herein are those of the author(s).

Motivation

- State-of-the-art: wide literature about the efficiency of secondary schools within countries ➔ lack of studies in an international perspective
  - Data comparability, policy implications

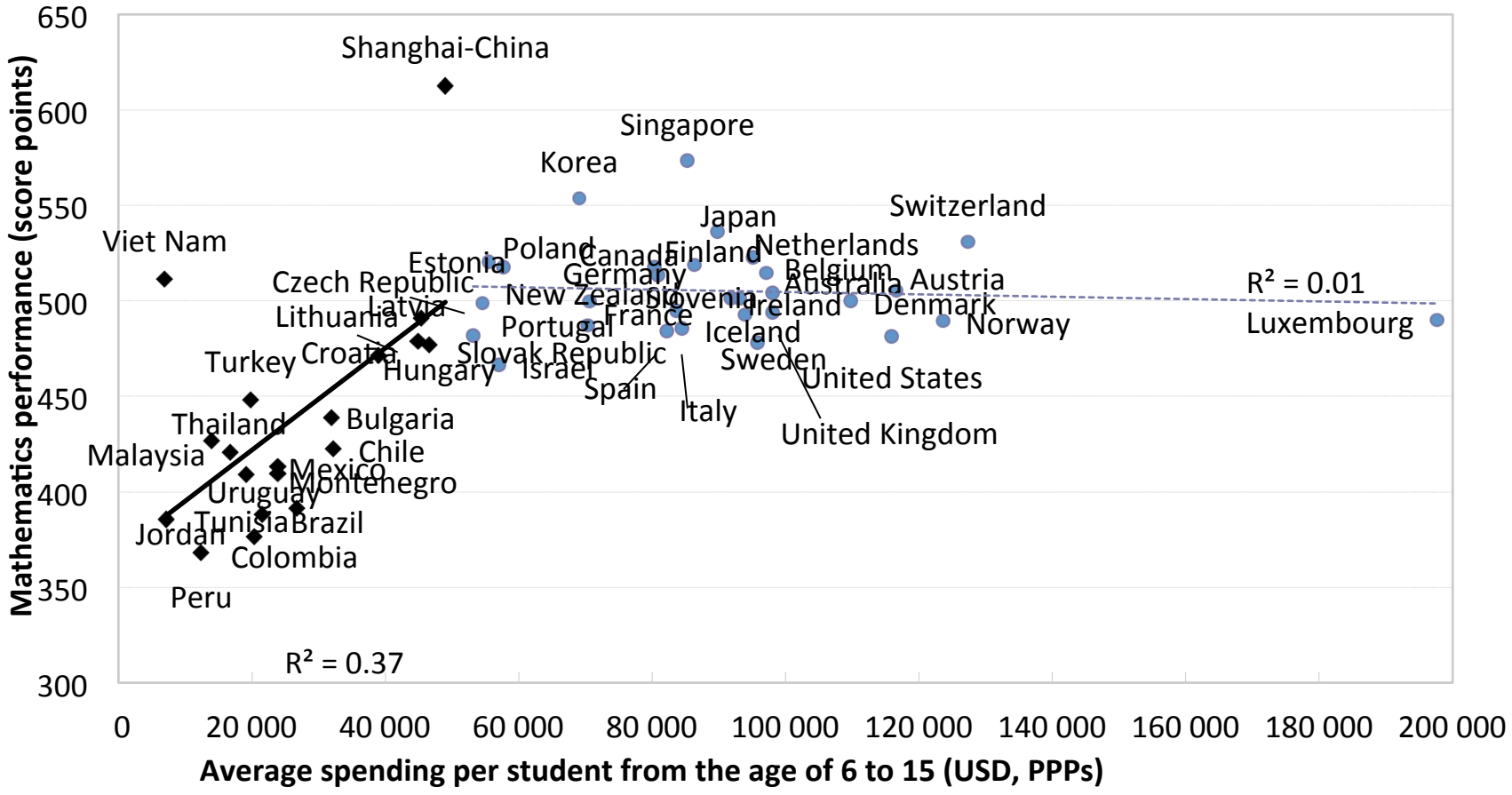
- From the comparison of countries' educational performances (i.e. test scores) to the idea of an international benchmark for schools' efficiency (outputs/inputs)
  - Use of OECD-PISA data

Research questions

- How does the efficiency of schools in the high-spending countries look like, in a comparative perspective?
  - Are the differences between countries higher or lower than those within countries?
- Which are the main factors associated with the efficiency of schools, when estimated using an international benchmark?
  - Heterogeneity of these factors?
- How is the (eventual) link between schools' efficiency and equity?
  - Equality and/or inclusion

Selection of countries

- High spending countries (PISA 2013, Fig. IV.I.8)



Empirical model

- Data Envelopment Analysis
  - Output orientation
  - Variable returns to scale (VRS)
  - Simar & Wilson (2000) bootstrap procedure (2,000 replicates) [estimated with Benchmarking © in R]

- Robustness check:
  - SFA
  - DEA with different combinations of inputs and outputs

Critical assumptions

- PISA is not explicitly designed for being used at the school-level
- Critical assumptions for conducting this study:
  - The students selected within the school are representative of those attending the school
  - The schools selected are representative at the country-level (or, they are representative of those schools that educate 15 y.o. students; ISCED1 schools are excluded)
  - The measurement of efficiency scores capture the contribution of the school net of the students' story until this moment ➔ the role of ESCS

Selection of inputs and outputs

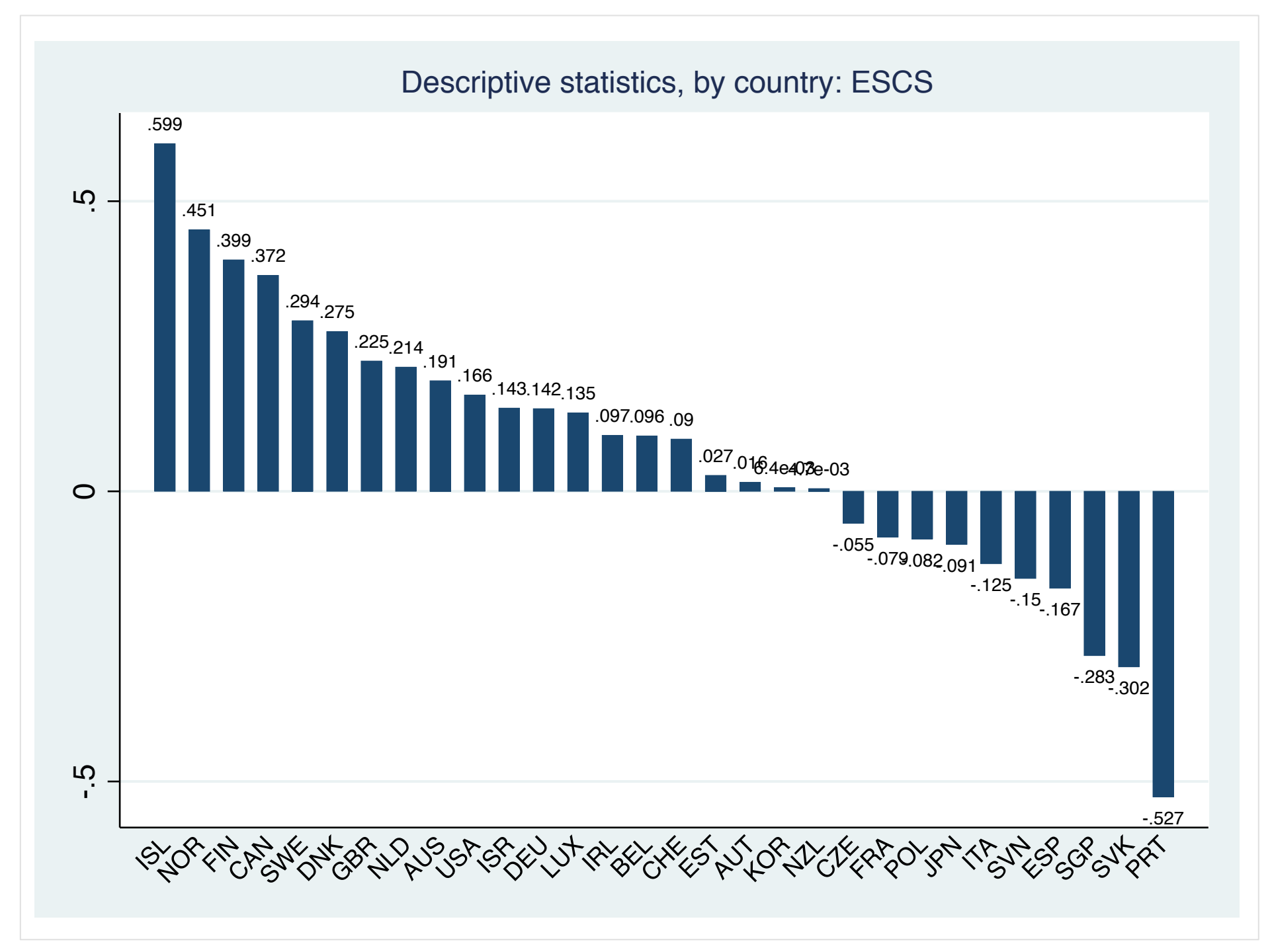
- Literature, conceptual production function:
  - Quality/quantity of human and financial resources
  - *Expenditures – missing*
  - Students' socioeconomic background
  - Cognitive skills (i.e. test scores)
  - *Success in education (pass rates, graduation, etc.) – missing*
  - *Non-cognitive skills – missing*

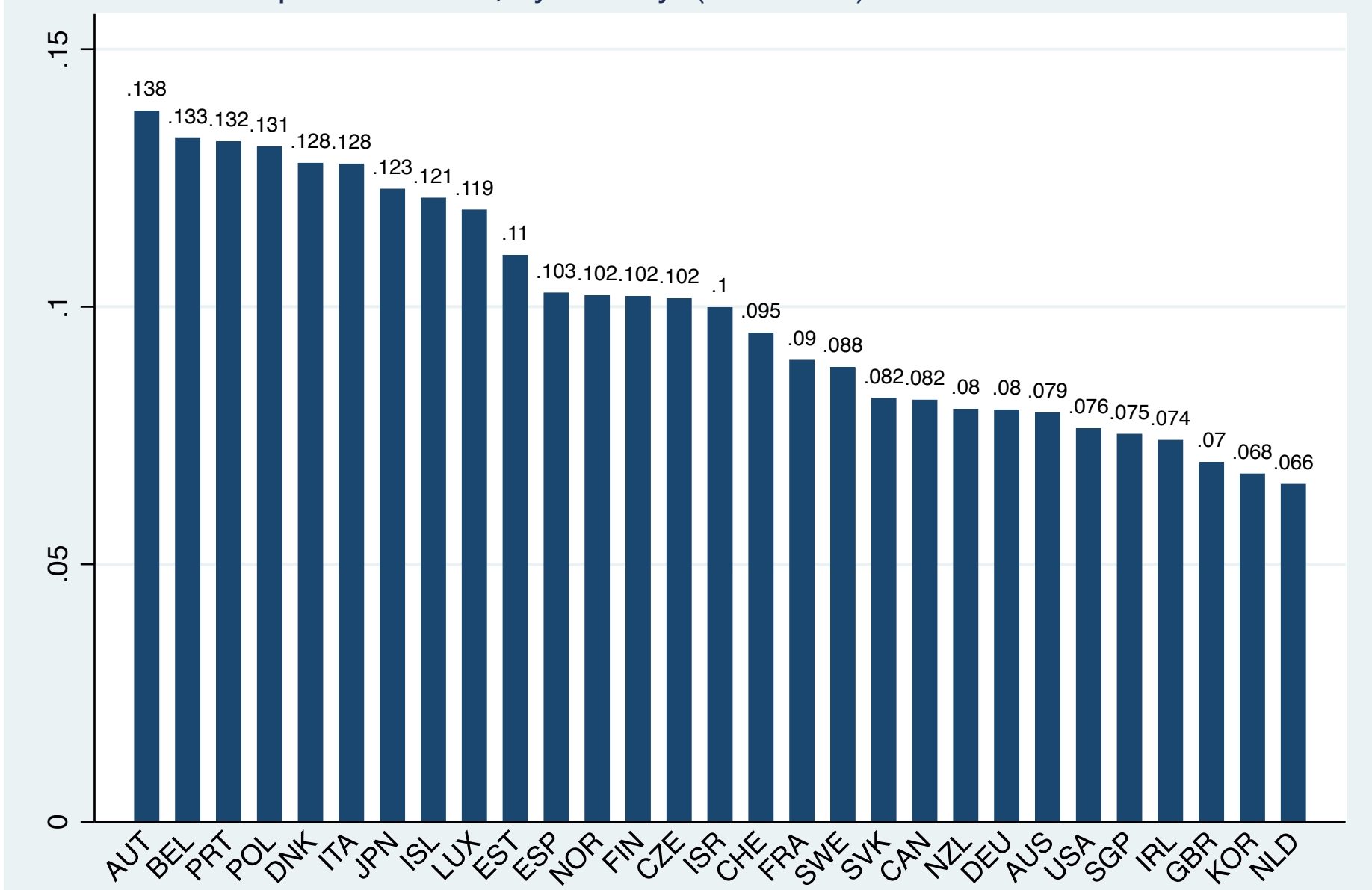
- Final selection
  - (Inverse of) students/teachers ratio (*St_Ratio*)
  - Number of computers per student (*Computer_n*)
  - Students' average *ESCS* (index of Economic, Social and Cultural Status)
  - Average test score in mathematics (*pv1math*) and reading (*pv1read*)

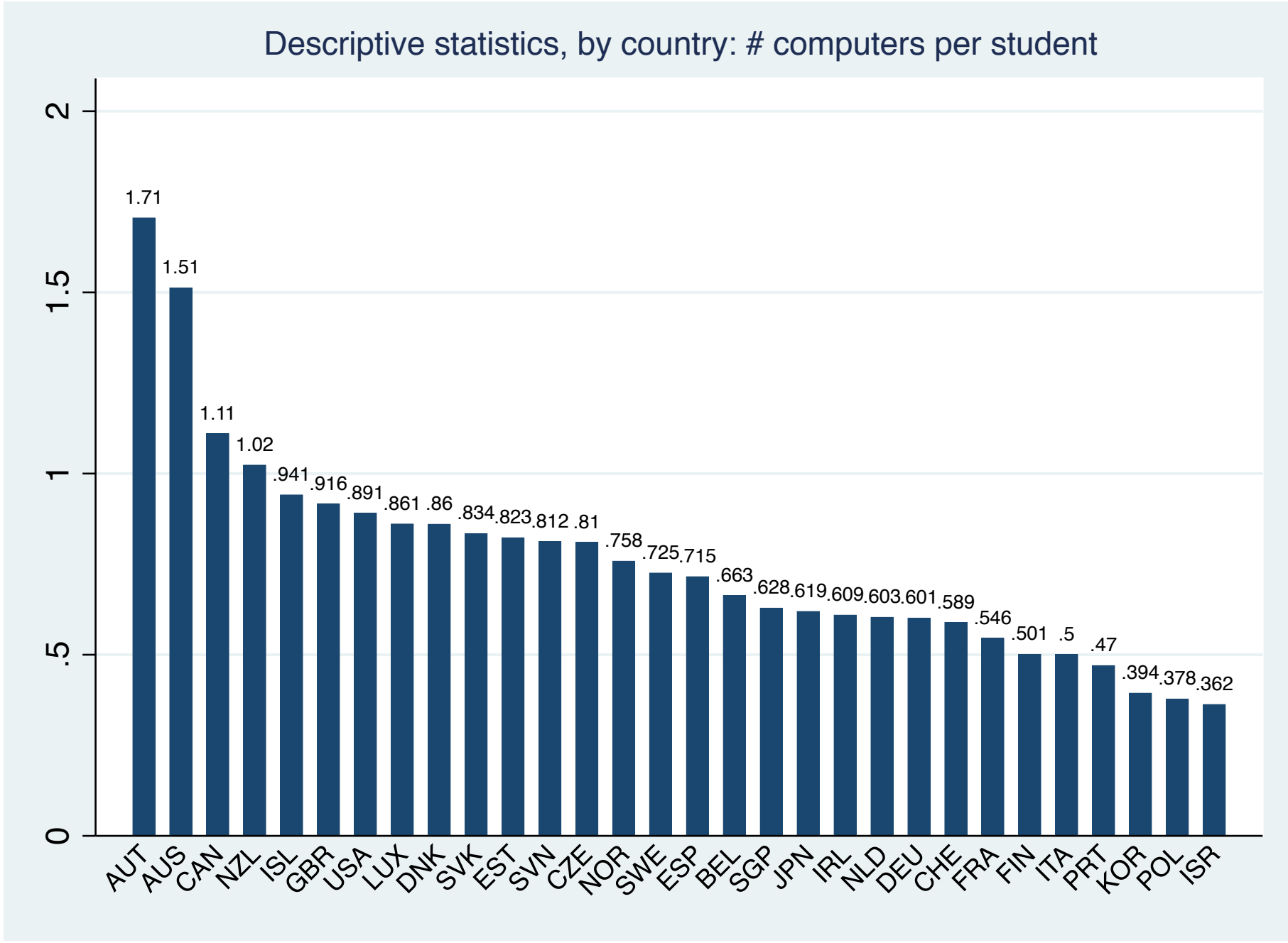
Inputs and outputs, descriptive statistics

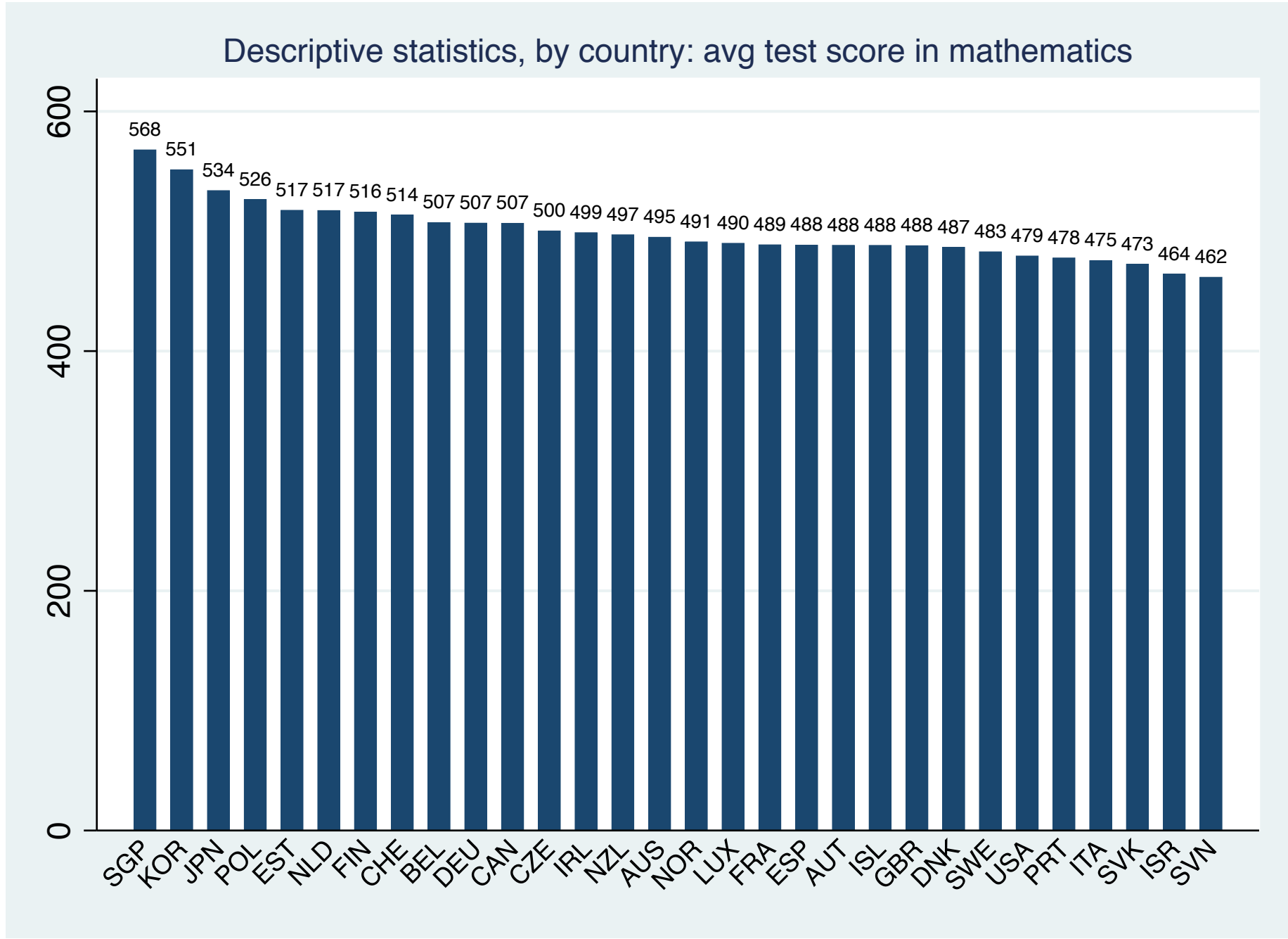
- In the overall sample: 8,640 schools

Variable	Mean	Std. Dev.	Min	Max
ESCS	0.063	0.528	-2.636	1.578
StRatio	0.105	0.141	0.007	5.988
Computer_n	0.836	1.250	0	55
pv1math	496.583	63.736	98.232	782.373
pv1read	495.393	66.350	163.594	734.684


Descriptive statistics, by country: ESCS
.5
0
-.5
.599
.451
.399
.372
.294
.275
.225
.214
.191
.166
.143
.142
.135
.097
.096
.09
.027
.016
-.055
-.079
-.082
-.091
-.125
-.15
-.167
-.283
-.302
-.527
ISL
NOR
FIN
CAN
SWE
DNK
GBR
NLD
AUS
USA
ISR
DEU
LUX
IRL
BEL
CHE
EST
AUT
KOR
NZL
CZE
FRA
POL
JPN
ITA
SVN
ESP
SGP
SVK
PRT


Descriptive statistics, by country: (inverse of) students/teachers ratio
0
.05
.1
.15
.138
.133
.132
.131
.128
.128
.123
.121
.119
.11
.103
.102
.102
.102
.1
.095
.09
.088
.082
.082
.08
.08
.079
.076
.075
.074
.07
.068
.066
AUT
BEL
PRT
POL
DNK
ITA
JPN
ISL
LUX
EST
ESP
NOR
FIN
CZE
ISR
CHE
FRA
SWE
SVK
CAN
NZL
DEU
AUS
USA
SGP
IRL
GBR
KOR
NLD

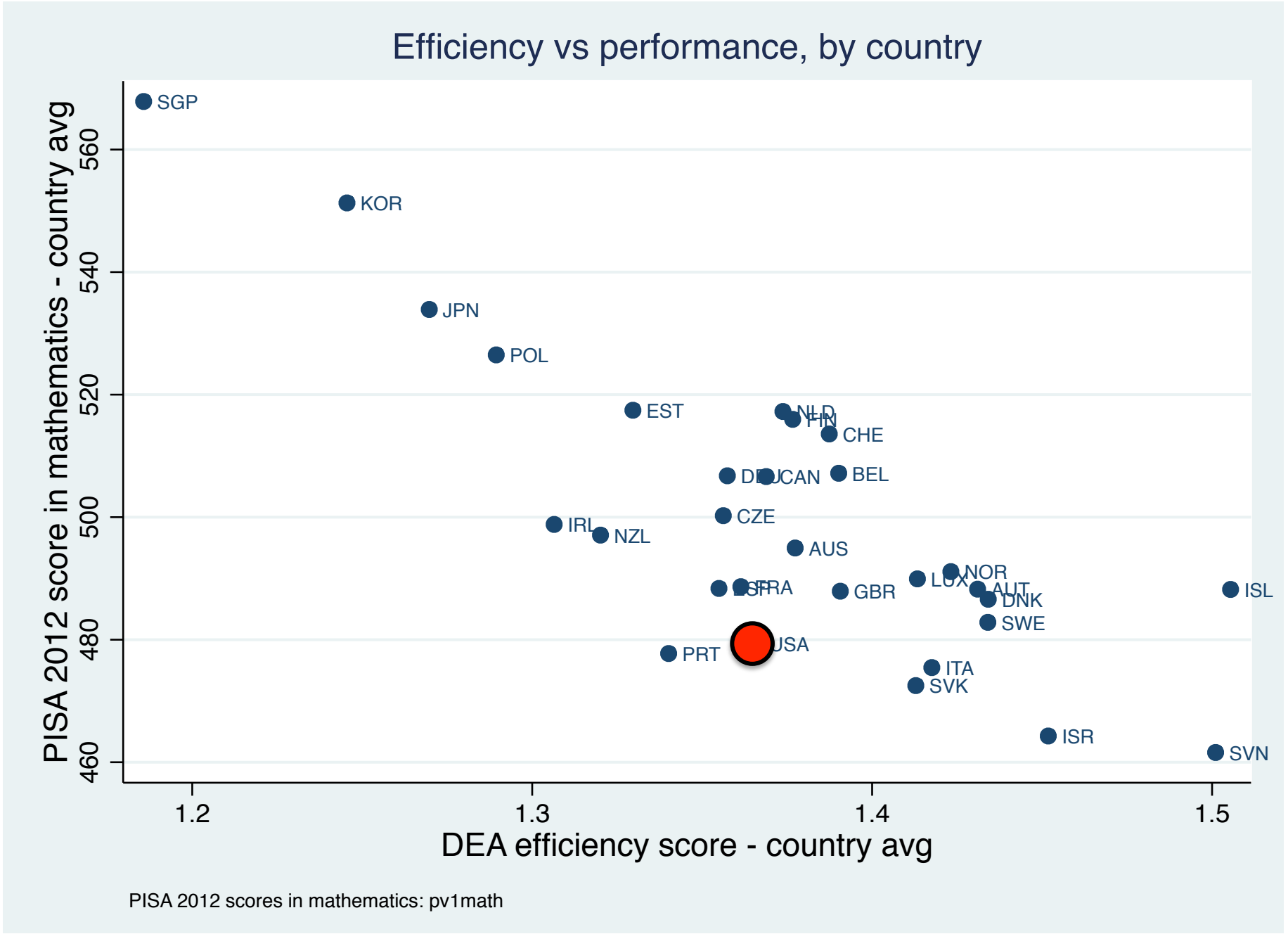

Descriptive statistics, by country: # computers per student
0
.5
1
1.5
2
1.71
1.51
1.11
1.02
.941
.916
.891
.861
.86
.834
.823
.812
.81
.758
.725
.715
.663
.628
.619
.609
.603
.601
.589
.546
.501
.5
.47
.394
.378
.362
AUT
AUS
CAN
NZL
ISL
GBR
USA
LUX
DNK
SVK
EST
SVN
CZE
NOR
SWE
ESP
BEL
SGP
JPN
IRL
NLD
DEU
CHE
FRA
FIN
ITA
PRT
KOR
POL
ISR


Descriptive statistics, by country: avg test score in mathematics
600
400
200
0
568
551
534
526
517
517
516
514
507
507
507
500
499
497
495
491
490
489
488
488
488
488
487
483
479
478
475
473
464
462
SGP
KOR
JPN
POL
EST
NLD
FIN
CHE
BEL
DEU
CAN
CZE
IRL
NZL
AUS
NOR
LUX
FRA
ESP
AUT
ISL
GBR
DNK
SWE
USA
PRT
ITA
SVK
ISR
SVN

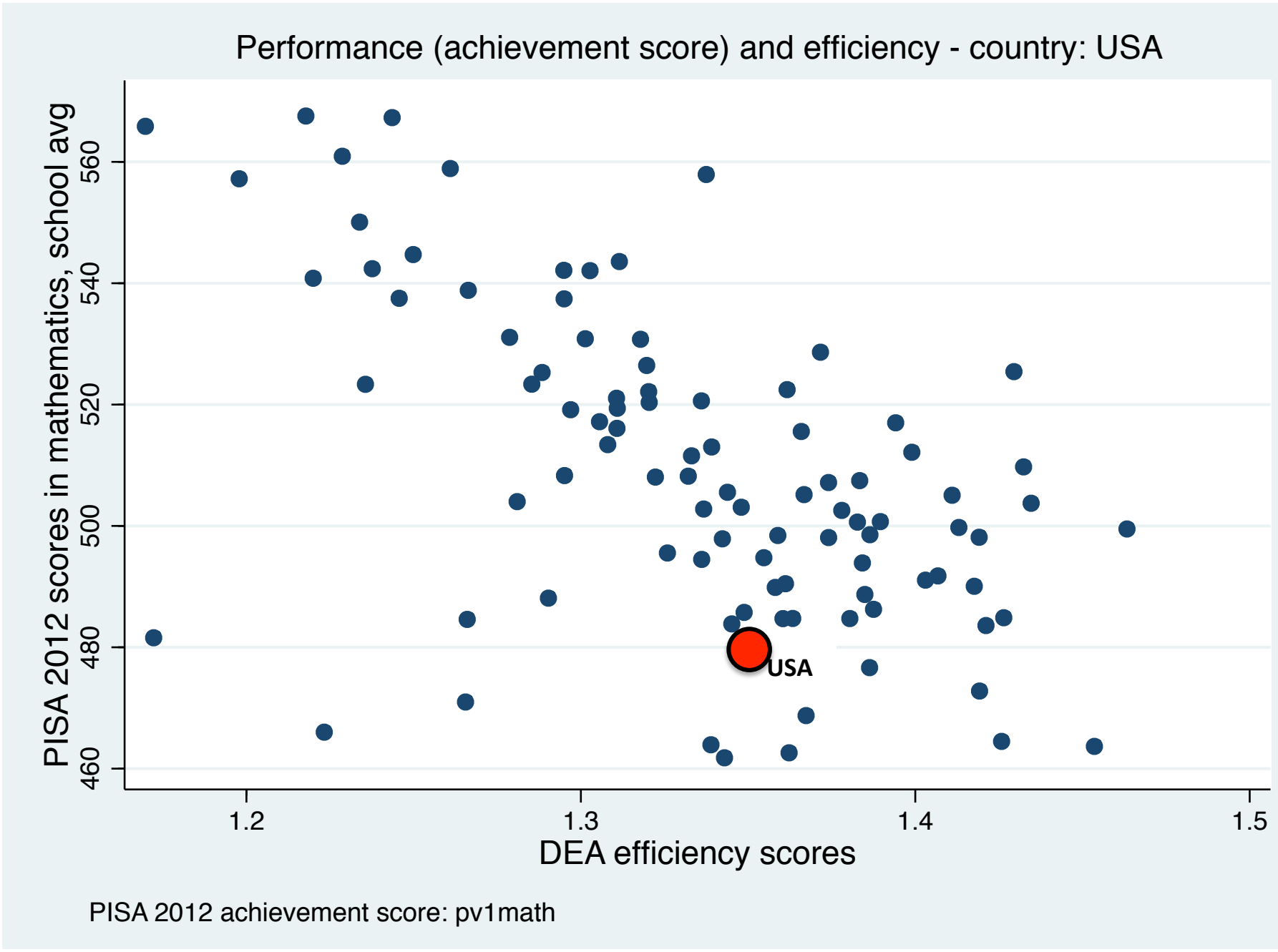
Results: efficiency scores

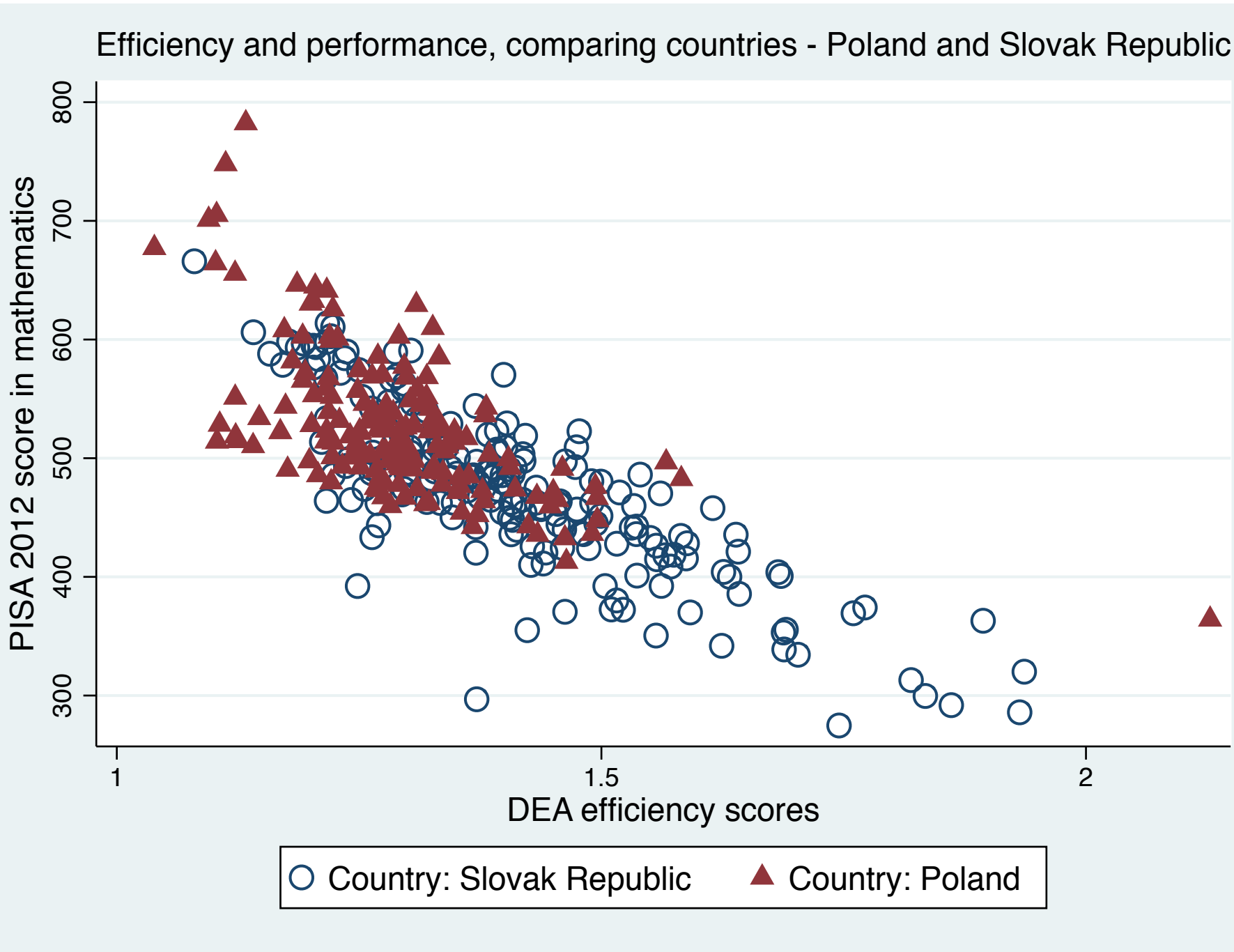
- It is interesting to show not only mean schools' efficiency scores, but their distribution by country
  - How much differentiated are the educational systems, *within*?
  - How many schools are more efficient than the average – as computed through the international benchmark?

- There is NOT the "average" Italian, Spanish, American... etc. school
  - the main message: efficiency is a property of schools, not countries – the focus is not on structural differences
  - for the single school: the international benchmark allows a wider set of efficient solutions to look at for improving its own activities



PISA 2012 scores in mathematics: pv1math

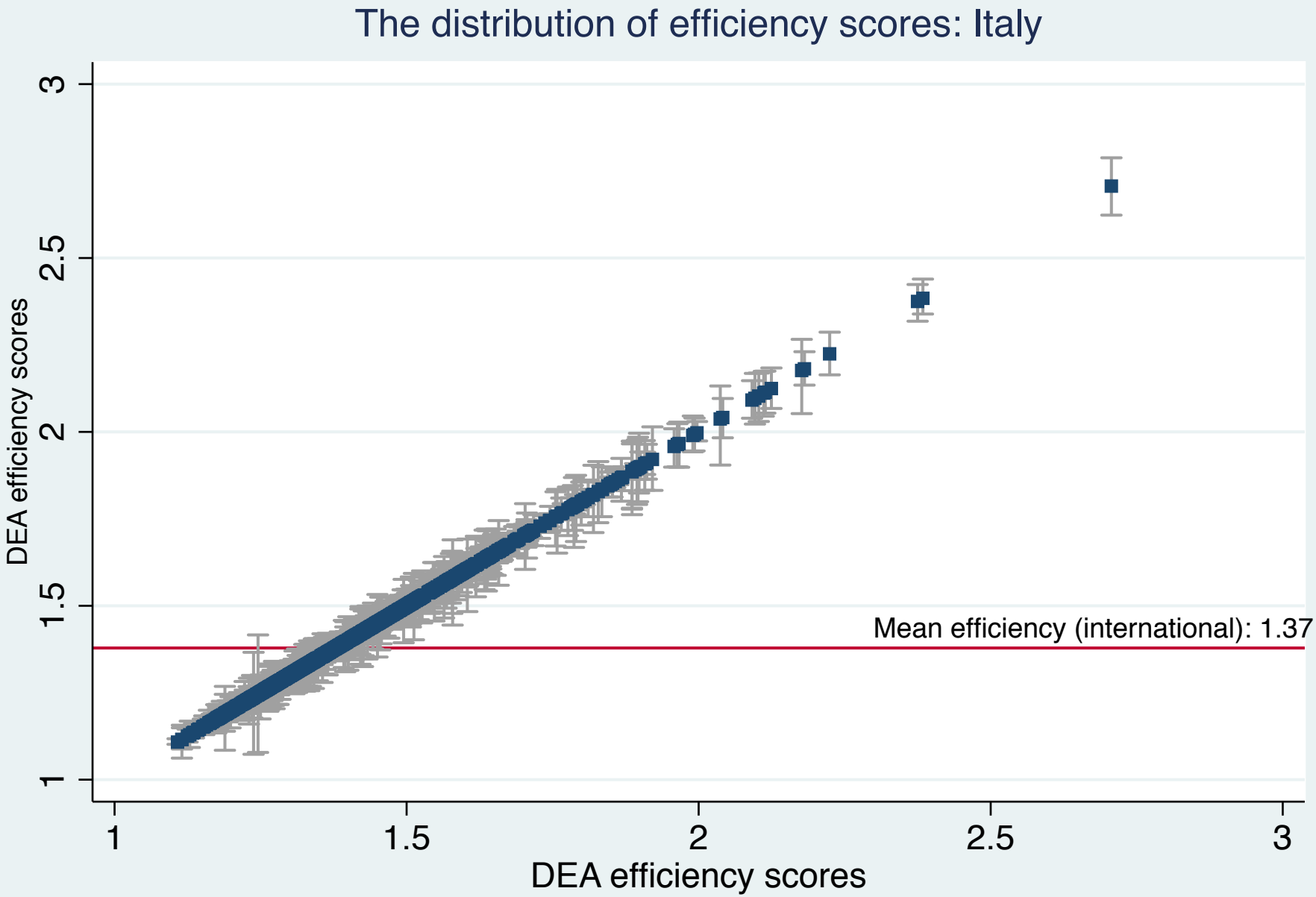

Performance (achievement score) and efficiency - country: USA
PISA 2012 scores in mathematics, school avg
460
480
500
520
540
560
USA
1.2
1.3
1.4
1.5
DEA efficiency scores
PISA 2012 achievement score: pv1math

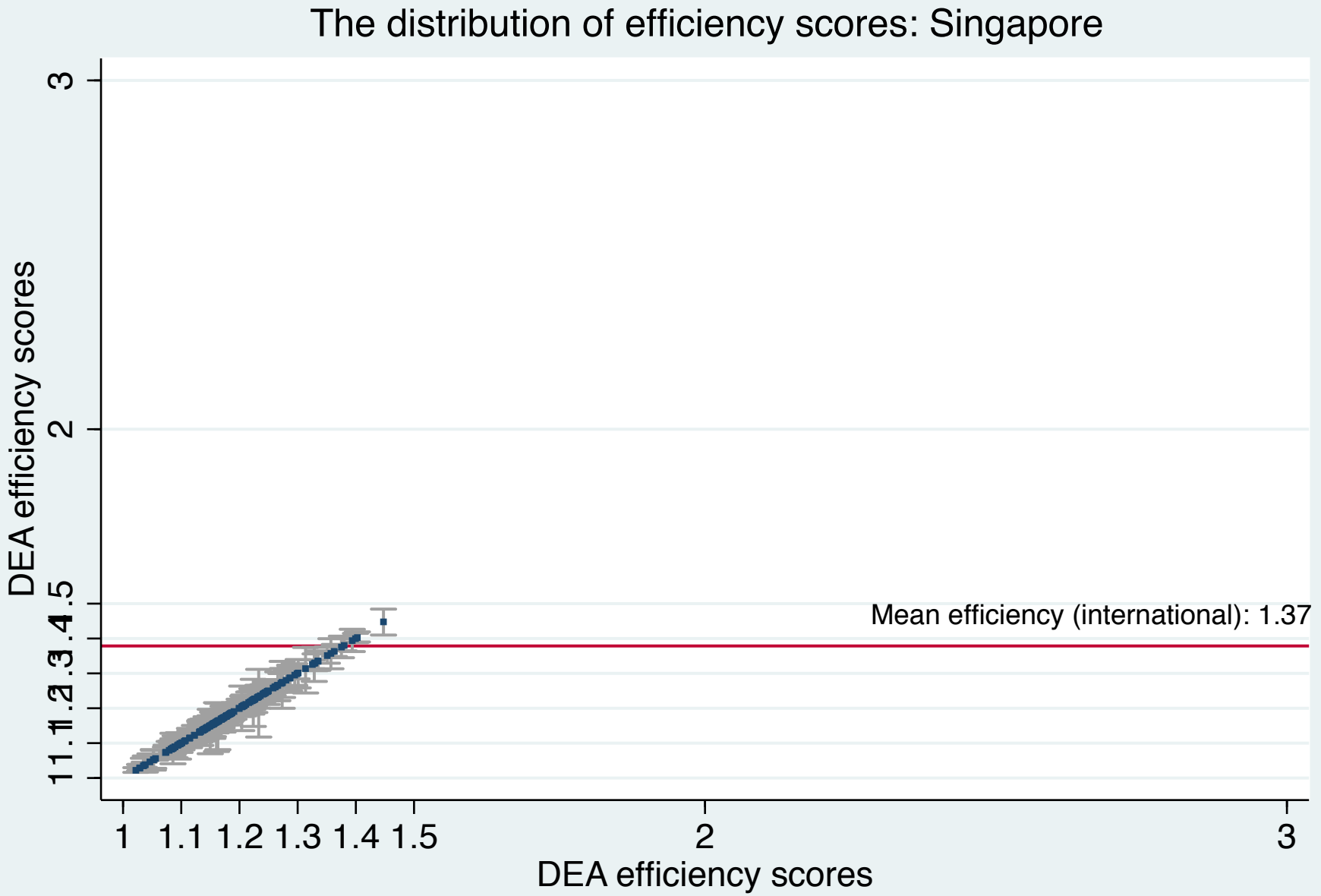

Efficiency and performance, comparing countries - Poland and Slovak Republic
PISA 2012 score in mathematics
300
400
500
600
700
800
1
1.5
2
DEA efficiency scores
Country: Slovak Republic
Country: Poland

Efficient and inefficient schools

- How many schools are more/less efficient than the average, by country? [The CI does not cross the average value]

		More efficient than avg		Less efficient than avg	
Selected countries	n	n	%	n	%
AUS	718	369	51%	290	40%
DEU	194	110	57%	64	33%
EST	199	128	64%	51	26%
IRL	**152**	**125**	**82%**	**18**	**12%**
ISR	**141**	**60**	**43%**	**69**	**49%**
ITA	1,044	488	47%	459	44%
JPN	190	143	75%	33	17%
PRT	171	99	58%	40	23%
SGP	**163**	**155**	**95%**	**3**	**2%**
USA	152	73	48%	41	27%
Total	8,640	4,377	51%	3,325	38%


The distribution of efficiency scores: Italy
DEA efficiency scores
1
1.5
2
2.5
3
Mean efficiency (international): 1.37
DEA efficiency scores

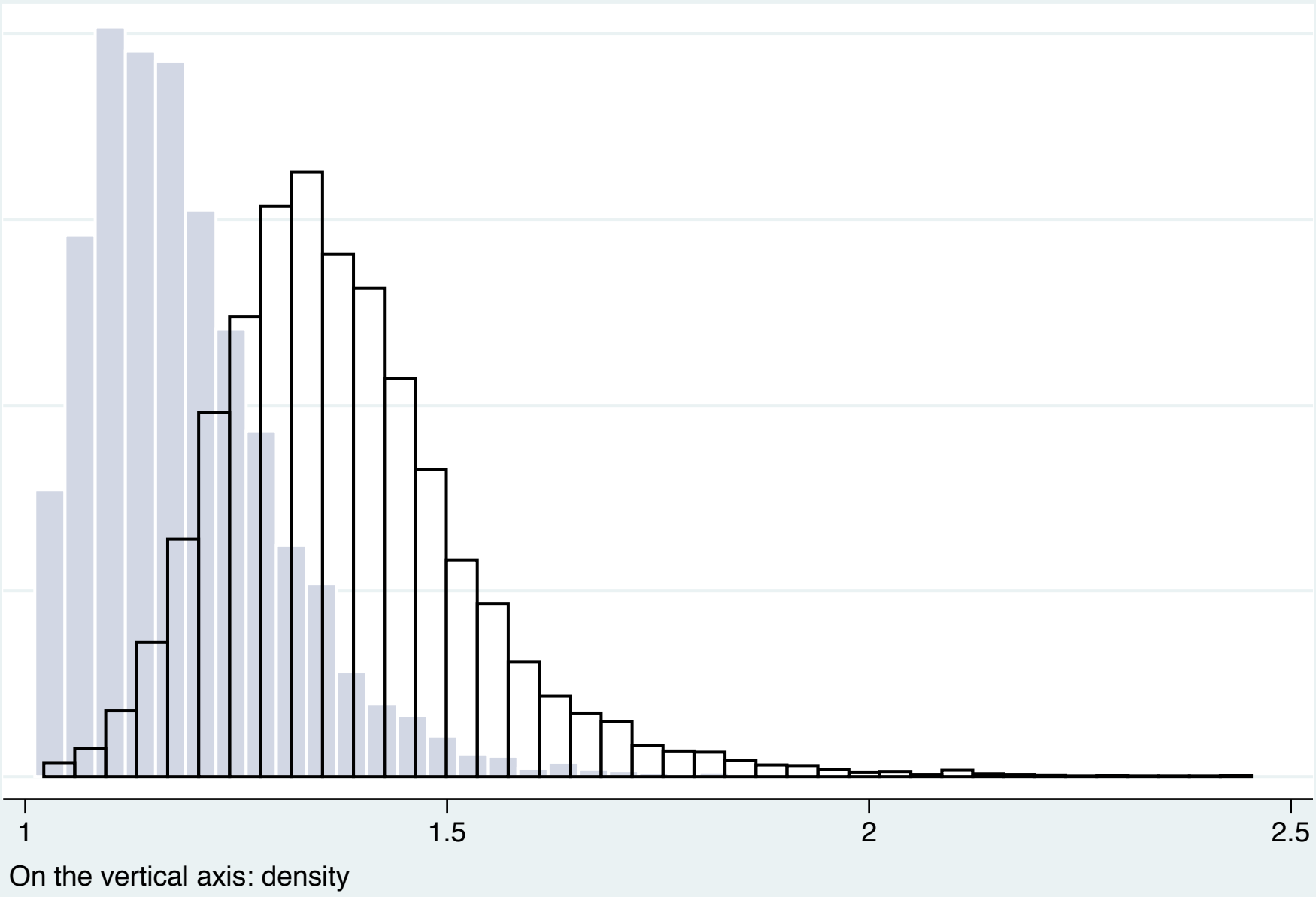

The distribution of efficiency scores: Singapore
DEA efficiency scores
1
1.1
1.2
1.3
1.4
1.5
2
3
Mean efficiency (international): 1.37
DEA efficiency scores

Country-specific frontiers

- DEA model has been re-estimated for each country, *separately*

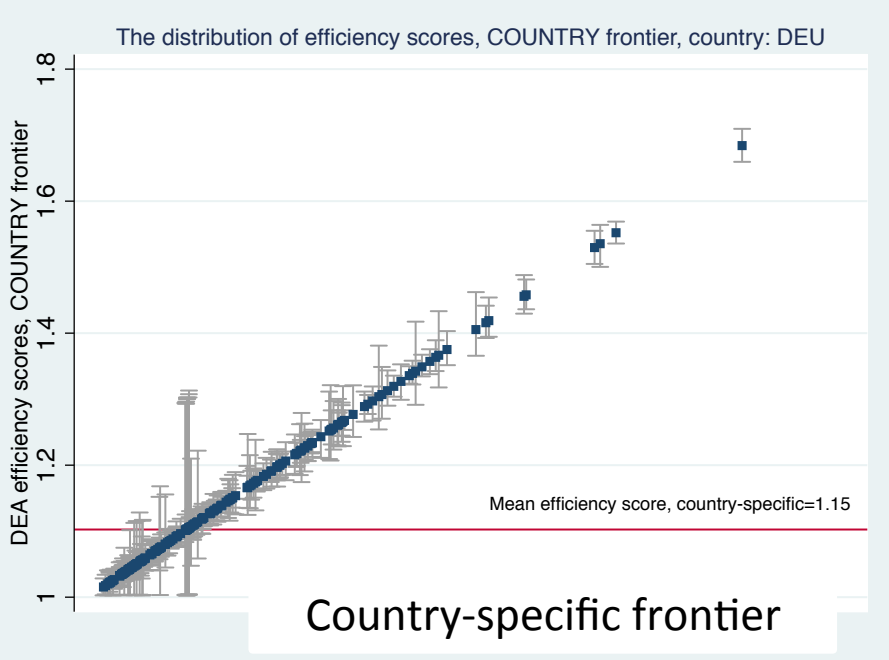
- If the differences between countries are bigger than those within countries (i.e. between schools within country), we should observe:
  - The estimated correlations between the two sets of efficiency scores (international frontier vs country-specific frontier) being very different;
  - The distribution of scores being different – in some or all countries

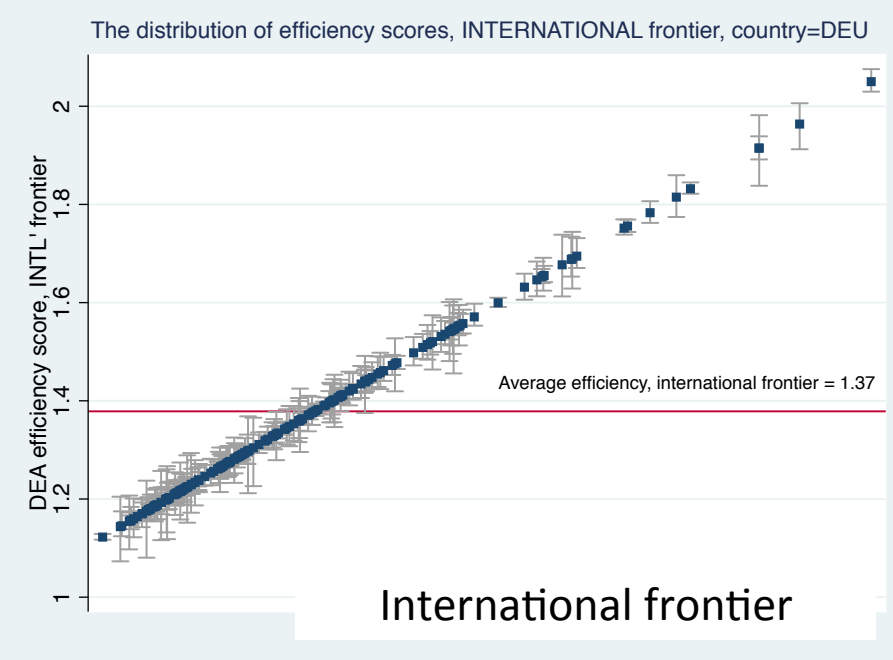
  - Important: a simple comparison of efficiency scores is meaningless!
  - Important (2): the confidence intervals are smaller when the international frontier is estimated as benchmark


The distribution of efficiency scores: international frontier (white) vs country-specific frontier (blue)
1
1.5
2
2.5
On the vertical axis: density

Country specific frontier: DEU

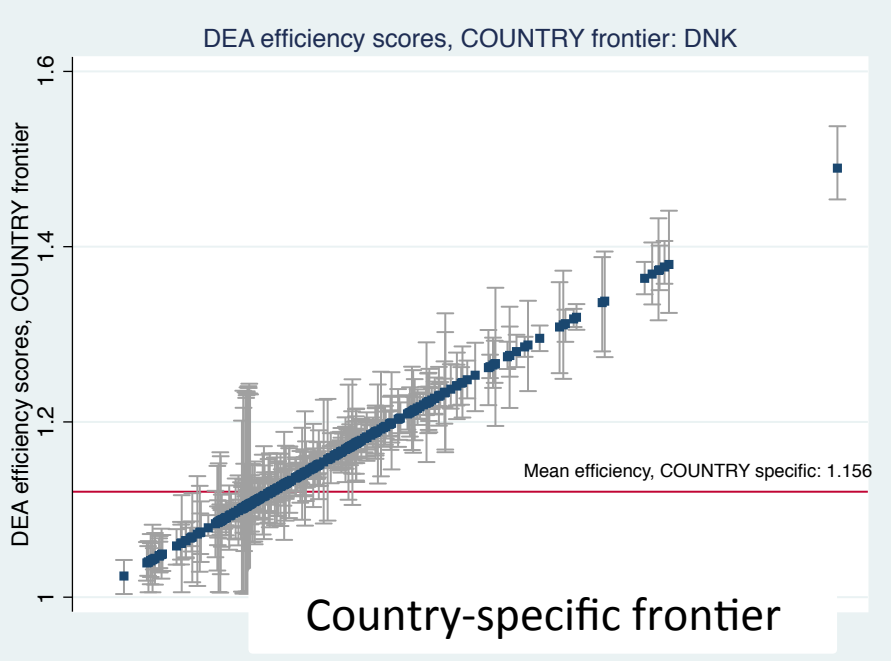
- For some countries, the comparison with an international benchmark implies a higher proportion of efficient schools – i.e. they are determining more frequently the set of efficient schools internationally
  - Example: GERMANY

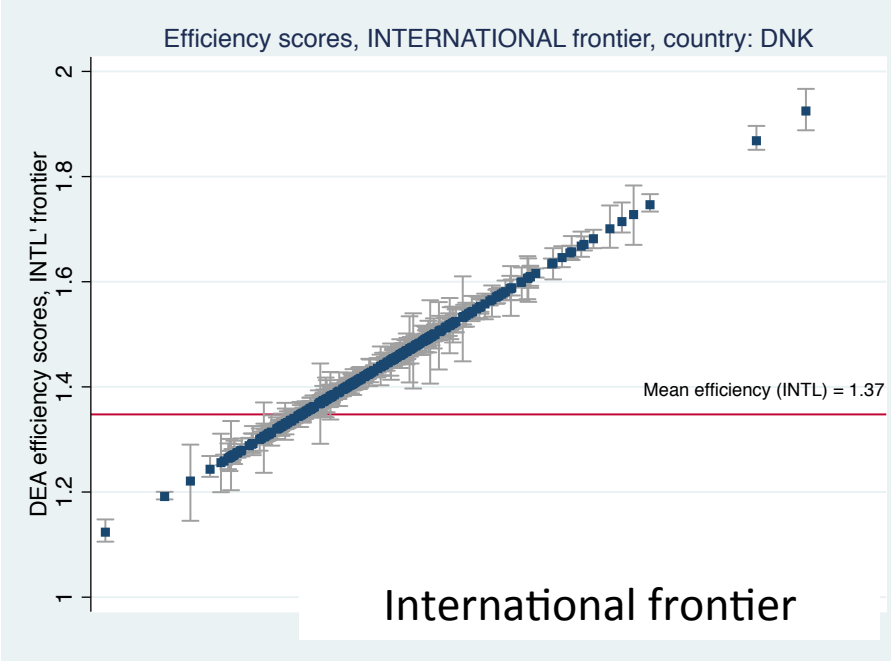




Country specific frontier: DNK

- For some countries, the comparison with an international benchmark implies a lower proportion of efficient schools – i.e. they are penalised by the comparison with the set of efficient schools internationally
  - Example: DENMARK





Country specific frontiers

- Some examples of countries for which the estimation of a country-specific frontier matters a lot (both negative/ positive effects)

		More efficient than avg, INTL frontier		More efficient than avg, COUNTRY frontier	
Country	n	n	%	n	%
CAN	753	391	52%	318	42%
CZE	249	144	58%	103	41%
DEU	194	110	57%	94	48%
DNK	**283**	**74**	**26%**	**106**	**37%**
EST	199	128	64%	82	41%
FRA	193	111	58%	83	43%
IRL	152	125	82%	61	40%
JPN	190	143	75%	72	38%

Factors associated w/efficiency scores (1)

- Second-stage Tobit regression
  - Dependent variable: bias-corrected DEA score
  - Backward and forward automatic procedure over a wide set of variables
  - Robustness check: Simar & Wilson (2007) – double-bootstrap procedure [done, results qualitatively and quantitatively similar]
- Set of variables
  - **School's general characteristics** (orientation; isced2, dispersion of scores and ESCS, private, size and class size)
  - **Students' characteristics** (% females, immigrants, repeaters, students who skipped school days; hours spent for homework)
  - **Schools' practices, resources and processes**
  - Program type and country dummies

Factors associated w/efficiency scores (2)

	Backward Coef.	se	Forward Coef.	se
School's general characteristic				
Program's orientation: general	-0.017***	[0.005]	-0.022***	[0.004]
% students below proficiency level 2	0.401***	[0.007]	0.398***	[0.007]
Private	0.033***	[0.003]	0.033***	[0.003]
Class size (avg) $<$ 15 students	0.014**	[0.005]	0.016**	[0.005]
Students' characteristic				
% immigrants 1st generation	-0.264***	[0.012]	-0.259***	[0.012]
% female students	-0.054***	[0.006]	-0.053***	[0.006]
Standard deviation of ESCS	0.148***	[0.007]	0.149***	[0.007]
% students who skipped 1 or + days	0.027*	[0.010]	0.031*	[0.009]
School's practices and processes				
Students report bad relations w/teachers	0.156***	[0.012]	0.158***	[0.012]
% certified teachers	-0.063***	[0.007]	-0.063***	[0.007]
Principal responsible - budget allocation	-0.009**	[0.003]	-0.008**	[0.003]
Principal organises teaching meetings	-0.015***	[0.003]	-0.015***	[0.003]
School organises volunteering	-0.028***	[0.003]	-0.029***	[0.003]
Constant	1.350	0.014	1.276	0.011

Factors associated w/efficiency scores (3)

- Exploring heterogeneity
  - **Across different countries**: running regressions separately and see which variables are more frequently associated with scores
  - **At different levels of the efficiency's distribution**: quantile regression

Factors associated w/efficiency scores (4)

	Statistically significant	Positively associated w/ eff	Negatively associated w/ eff
School's general characteristics			
Program's orientation: general	8	6	2
Private	14	3	11
Class size (avg) $<$ 15 students	12	4	8
Students' characteristics			
Standard deviation of ESCS	**14**	**6**	**8**
% female students	17	17	0
Hours spent for homework	**11**	**9**	**2**
% students who skipped 1 or + days	17	5	12
School's practices and processes			
Principal responsible - budget allocation	9	8	1
Index of teachers participation/governance	7	4	3
Achievement scores publicly available	9	3	6
School organises volunteering	8	5	3
Schools practices	8	2	6
School competes for students w/2 schools or +	10	4	6

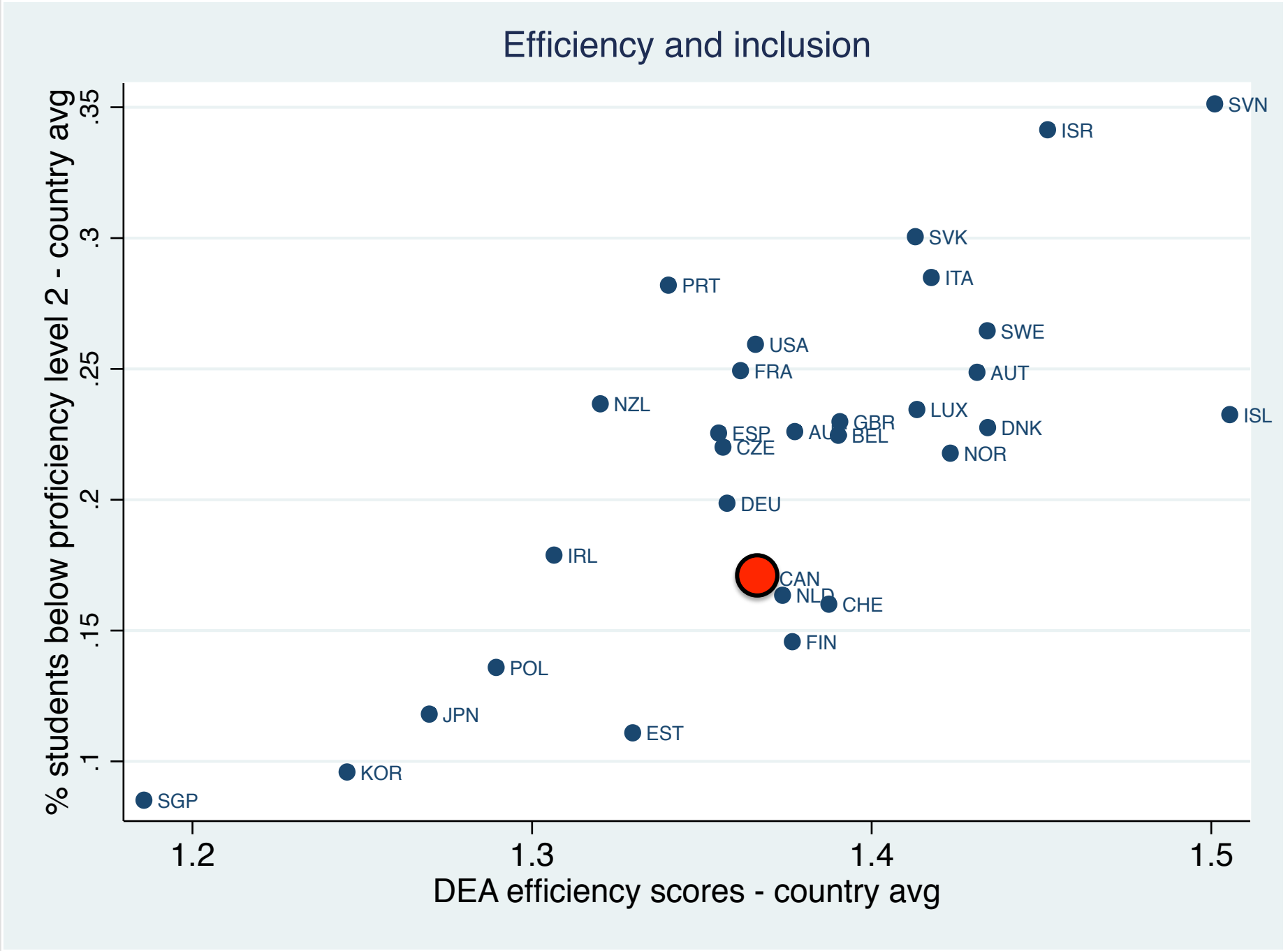
Factors associated w/efficiency scores (5)

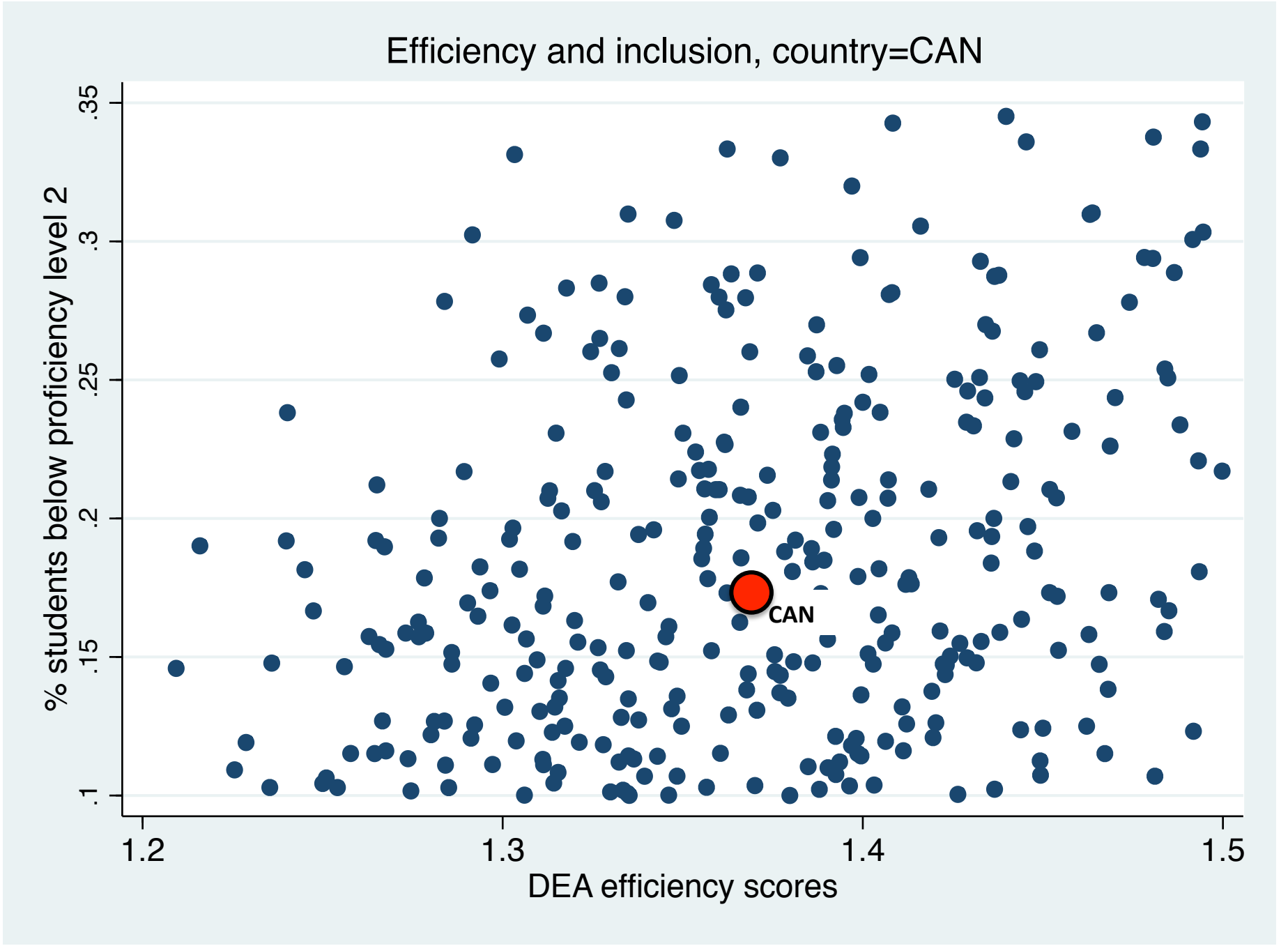
	25th percentile	Median	75th percentile
Program's orientation: general	-0.03827***	-0.02618***	-0.01255*
	0.009	*0.006*	*0.006*
Private school	0.00950*	0.03250***	**0.03094*** **
	0.005	*0.004*	*0.003*
Class size (avg) $<$ 15 students	**0.03189*** **	-0.01768***	-0.00192
	0.006	*0.005*	*0.005*
% students immigrants 1st	-0.45402***	-0.21912***	-0.12258***
	0.017	*0.013*	*0.012*
% students who skipped 1 or + days	0.02547	0.03466***	0.04923***
	0.014	*0.010*	*0.010*
Achievement scores publicly available	0.01671***	0.01109***	0.00366
	0.004	*0.003*	*0.003*
Teachers are monitored by principal	**-0.02394*** **	-0.01633***	-0.00543*
	0.004	*0.003*	*0.003*
School organises volunteering	**-0.03205*** **	-0.02544***	-0.00385
	0.004	*0.003*	*0.003*

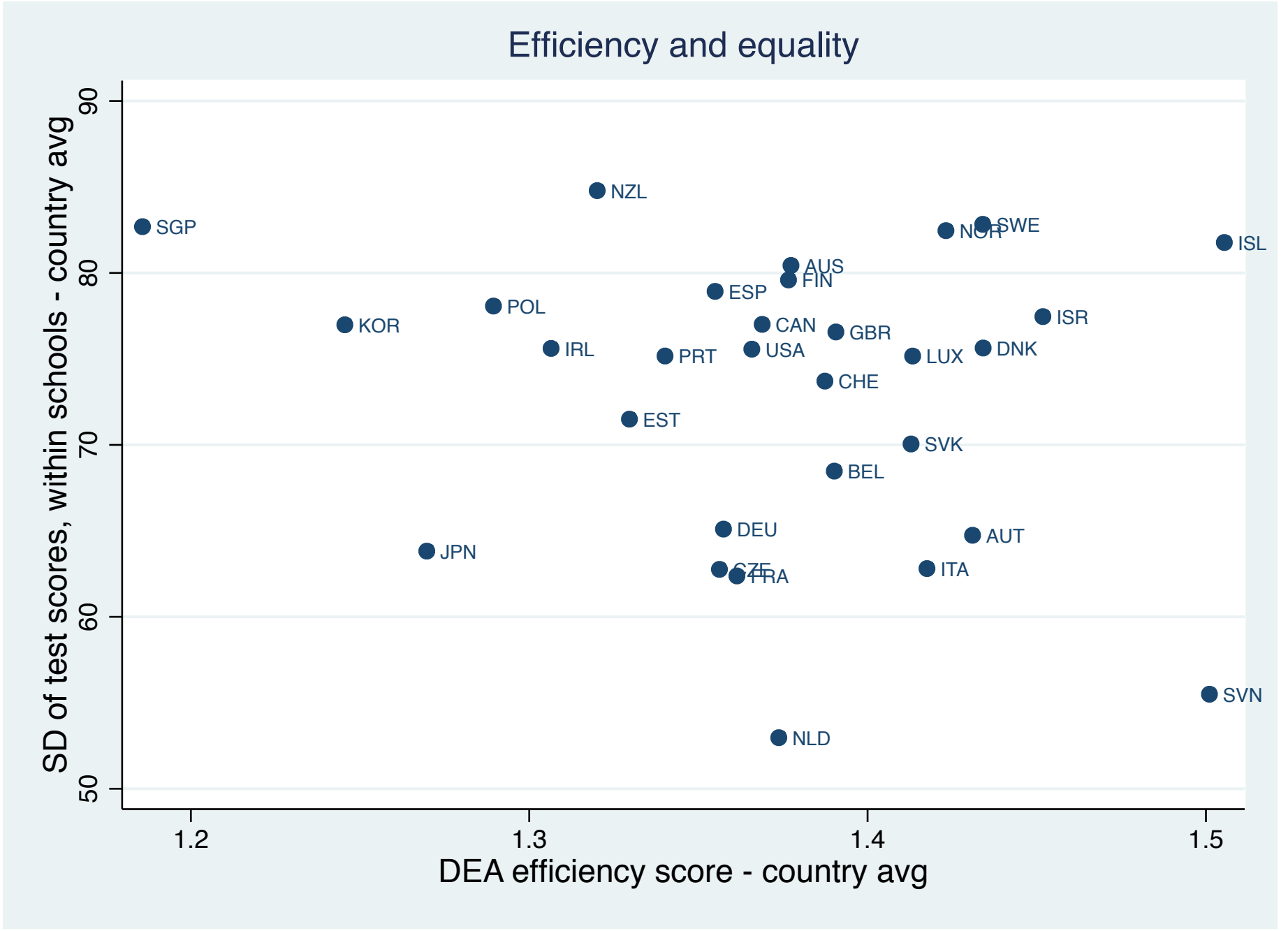
Efficiency and equity: correlations

- Equity definition(s)
  - **Inclusion**: the proportion of students below proficiency level 2 (Schleicher, 2014)
  - **Equality**: dispersion of scores around the mean (standard deviation), similar to the concept proposed by Freeman *et al.* (2010)

- Inclusion:
  - The "automatic" relationship with performance/efficiency (the lower the proportion of low performing students, the higher the performance/efficiency, all else equal) but (i) exceptions and (ii) different possibilities (i.e. higher equality for lower means)


Efficiency and inclusion
% students below proficiency level 2 - country avg
.1
.15
.2
.25
.3
.35
DEA efficiency scores - country avg
1.2
1.3
1.4
1.5
SVN
ISR
SVK
ITA
PRT
SWE
USA
FRA
AUT
NZL
LUX
ISL
GBR
DNK
ESP
AUS
BEL
CZE
NOR
DEU
IRL
CAN
NLD
CHE
FIN
POL
JPN
EST
KOR
SGP


Efficiency and inclusion, country=CAN
% students below proficiency level 2
.1
.15
.2
.25
.3
.35
1.2
1.3
1.4
1.5
DEA efficiency scores
CAN


Efficiency and equality
SD of test scores, within schools - country avg
50
60
70
80
90
DEA efficiency score - country avg
1.2
1.3
1.4
1.5
NZL
SGP
NOR
SWE
ISL
AUS
FIN
ESP
POL
KOR
CAN
GBR
ISR
IRL
PRT
USA
LUX
DNK
CHE
EST
SVK
BEL
DEU
AUT
JPN
CZE
FRA
ITA
SVN
NLD

Concluding remarks (1)

- Setting an international frontier for estimating schools' efficiency can be interesting if accepting the existence of comparable outputs and inputs ➔ a higher number (and types) of schools to be compared with
  - How taking the structural differences between countries into account? The related literature about "institutional settings" (Hanushek & Woessmann, 2010)

- Within-country differences are wider than between-countries (structural) ones
  - Does it make sense comparing the efficiency of countries' educational systems *as a whole*?

Concluding remarks (2)

- Potential extensions
  - New concepts and measures of equity ➔ for example, proportion of resilient students
  - How the distribution of efficiency scores (within and between countries) evolved over time? ➔ Comparing PISA 2003 and PISA 2012
  - Including some variable at country level for "explaining" structural differences in schools' efficiency across countries

- Adding further robustness checks:
  - SFA – Stochastic Frontier Analysis (Annex)
  - Different DEA specifications (*to be done*)
  - Eliminating outliers (*to be done*)

Annexes and additional materials

1. Robustness check: results from the SFA specification
2. Robustness check: results from alternative DEA specifications

Robustness check: results from SFA

- Stochastic Frontier Analysis (SFA)
  - Mathematics and Reading scores as alternative outputs

- Mathematical formulation (translog)

$$\ln(Math) = \beta_0 + \sum_i \beta_i lnX_i + \sum_i \beta_i lnX_i{}^2 + \frac{1}{2}\sum_i \sum_j \beta_i \beta_j \, lnX_i lnX_j + \delta_c + \varepsilon$$

  - Inputs as specified in the DEA formulation

SFA – correlation indexes

Pearson's correlation index	DEA	SFA (Math)	SFA (Reading)
DEA	1		
SFA (Math)	-0.7668*	1	
SFA (Reading)	-0.8361*	0.7965*	1
Spearman's correlation index	**DEA**	**SFA (Math)**	**SFA (Reading)**
DEA	1		
SFA (Math)	-0.6863*	1	
SFA (Reading)	-0.7862*	0.7881*	1