ETM 540 Group Project

Samira Akther
Jordan Hilton
Andey Nunes
Aparna Gandikota
Mohammed Sheikh

March 20, 2019

Abstract

The paper aims to study the efficiency of education systems in Organization for Economic Development (OECD) countries. We examine efficiency by looking at the relationship between defined inputs and outputs. Data Envelopment Analysis (DEA), which is a widely used tool for measuring efficiency, was conducted to measure the efficiency of the education system of the countries.

This project is available on GitHub: https://github.com/AndeyNunes/educationDEA

Introduction

In the last few decades, Benchmarking, a non-parametric evaluation approach, has found applications in a variety of fields. The technique has equipped organizations with the means to evaluate their processes and compare it to the best practice of peer organizations. This has allowed them to gauge their own performance and enable them to learn and develop plans to improve aspects of their own performance and set future targets [9].

In the education sector, studies utilizing benchmarking techniques are numerous and vary widely in focus. The richness of the data collected on education far exceeds that which have been thoroughly studied and published. This makes comparative studies on education data an interesting endeavor as it could reveal interesting patterns and discussion points.

In this paper, our aim is to study the efficiency of the education systems of 25 countries from the Organization for Economic Development (OECD) using DEA technique. Our reasons for working with OECD countries for this study include:

- All OECD countries dedicate a percentage of their Gross Domestic Product (GDP) to education spending and all participate in the Programme for International Student Assement (PISA).
- OECD collects data on several aspects relating to education such as spending, enrollment, graduation rates, etc. . . and publicly provides such data to any interested parties.
- The raw data can be considered to be reliable as both PISA and OECD are independently run organizations with a track record of professionalism and trustworthy in reporting accurate data.

Benchmarking with Data Envelopment Analysis (DEA)

There are few benchmarking technique currently in use including stochastic frontier analyis (SFA), but for purposes of this project, we will heavliy rely on Data Envelopment Analysis (DEA). DEA is a widely used benchmarking technique that was first proposed by Charnes, Cooper, and Rhodes in 1978. The technique is a nonparametric linear programming method that compares relative efficiencies of Decision Making Units (DMUs) and provides a comprehensive measure of each DMU's performance. In a DEA analysis, efficiency is usually given as the weighted sum of outputs dividied by the weighted sum of the input resources [2,3,6].

$$\text{Efficiency} = \text{Weighted} \sum \frac{outputs}{inputs}$$

DEA models can be categorized as either input oriented model or as output oriented model. Input-oriented models minimize inputs while satisfying at least the given output levels, while output oriented models maximize outputs without requiring more of any observed input values [7]. DEA models can also be categorized according to the weight constraints. In this case, Models are subdivided into four categories. Contant returns to scale CRS, where the CCR model is built upon and has no constraints. Variable returns to scale VRS where the sum of all lambda is set to 1. Increasing returns to scale IRS where is the sum is set to greater than 1 and Decreasing returns to scale DRS having the sum being less than 1 [2].

Returns to Scale	Envelopment Constraint
CRS	No constraint needed
VRS	$\sum_{j=1}^{N^D} \lambda_j = 1$
IRS	$\sum_{j=1}^{N^D} \lambda_j \geq 1$
DRS	$\sum_{j=1}^{N^D} \lambda_j \leq 1$

In this paper, our aim is to utilize output oriented variable returns to scale model to look at the efficiences of the 25 OECD countries in terms of spending as the input and the PISA test score as our output. Our objectives is to gauge how does spending translate to the performance of students. How close or far is each country's education system operating relative to the DEA frontier.

Literature Review

Data Envelopment Analysis has been used to evaluate educational efficiencies in OECD countries. [5]

Methodology

Benchmarking using a DEA model involves measuing the relative efficiency of a Decision Making Unit (DMU), which is determined by comparing the studied unit's performance in producing output with a target. [2] In our study, DMUs are the country's education system. The objective here is to find the efficiency of this education system in making educational achievements, measured by the PISA test scores and college graduation rates, using the resources invested in the system, measured in time and money.

For inputs, we considered each country's education spending per student in US dollars and the number of hours teachers spend with students. Per student data will help us to eliminate the size effect of a country with large GDP and thus large budget for education.

As our goal is to evaluate the efficiency of the country's education system, we decided to use PISA score as measumement of success. PISA is widely recognized performance yardstick to assess the competency of a country's school system. PISA focus on three segment: reading, mathematic and science.

Our project looks first at a single input of Total Spending per student and single output of average PISA score. We then expand on the model by including additional input of Total Hours teachers spend along with the more granualar PISA scores and an additional output, tertiary graduation rate, for a multiple input - multiple output model.

In each case, our DEA model is based on output-orientation because the focus of the model is on improving and optimizing the output based on the input as observed in studied units. We assumed that each country

has a given fund for spending on education which is influened by various factors including the GDP, education budgets, and political factors.

The basic DEA model is based on an assumption of constant returns to scale. This means that efficient output with a given input can be scaled up or down in a linear or constant rate. Anderson describes this as exhibiting no limit to how big an operation can get with the ratio of input, whereas often in reality, there are structural differnces that inhibit such behavior. [2] The DEA model in our study is based on variable returns to scale (VRS) on the nature of the inputs and outputs. With VRS specified, input can scale up or down, but the output may not scale up or down by the same amount. [3] In fact in the case of test scores, there is an upper limit on the output.

Consider n number of education system DMU_j which consumes $x_{i,j}$ inputs and produce $y_{r,j}$ output. The education system operates in variable returns to scale. So this output oriented model can be expressed by a linear algebric model where the objective will be to maximize the efficiency scores of j education system. The efficiency scores can be defined by ϕ

$$\text{Max } \phi$$
$$\text{subject to } \sum_{j=1}^{n} x_{i,j}\lambda_j \leq x_{i,k} \ \forall \ i$$
$$\sum_{j=1}^{n} y_{r,j}\lambda_j \geq \phi_{r,k} \ \forall \ r$$
$$\lambda_j \geq 0 \ \forall \ j$$
$$\sum_{j=1}^{n} \lambda_j = 1$$

Here, a country's education system is considered efficient if the efficeincy score ϕ is 1. The country with score of 1 lies in the efficient frontier which is a linear combination of all the countries with efficiency scores of 1. The countries with scores less than 1 will be considered as relatively inefficient in comparison to the target specified by λ. The vector λ is specific amount of a unit j used in setting the target for performance for studied unit k. To accomodate the variable returns to scale of our studied input and output, a constarint that λ sums up to 1 is added to the model. Also, as it is a output oriented model, the first (input) constraint will be satisfied while trying to exceed the second constraint (output) by as much possible.

The discussed methodology has been used to perform two DEA model and analysis for the 25 OECD countries for which we have all the data points.

The following OECD countries were dropped from the analysis due to missing data:

Data was obtained from secondary sources. Data on PISA scores and teaching hours were obtained from OECD website. Data on public education spending per student and tertiary graduation rate were obtained from UNESCO website.

Summary of the data are included in the Appendix.

DEA Model

Part A: Single input, single output

For this first part, we will set up and solve a DEA model using total educational spending per student for the input and the aggregated PISA test score average for the single output. The input output diagram for this model is included in the Appendix.

Table 2: Countries not included due to missing data

| Var1	Freq
	1
Australia	1
Belgium	1
Denmark	1
Estonia	1
France	1
Iceland	1
Ireland	1
Lithuania	1
Sweden	1
Switzerland	1
United Kingdom	1

```r
# Run a DEA with a single input (total spending) and a single output (Average PISA)
x <- DEAdata %>% select(Total.Spending) # input
y <- DEAdata %>% select(AveragePISA) # output

row.names(x) <- DEAdata$Countries # input labels
row.names(y) <- DEAdata$Countries # output labels

# output oriented model with variable returns to scale
ressingle <- DeaMultiplierModel(x, y, rts = "vrs", orientation = "output")
```

```r
dea.plot(x = x, y = y, txt = dimnames(x)[[1]], GRID = T,
         xlab = "Total Spending", ylab = "Average PISA",
         xlim = c(10000,120000),
         ylim = c(400,550), main = "2015 OECD data DEA plot")
```

2015 OECD data DEA plot
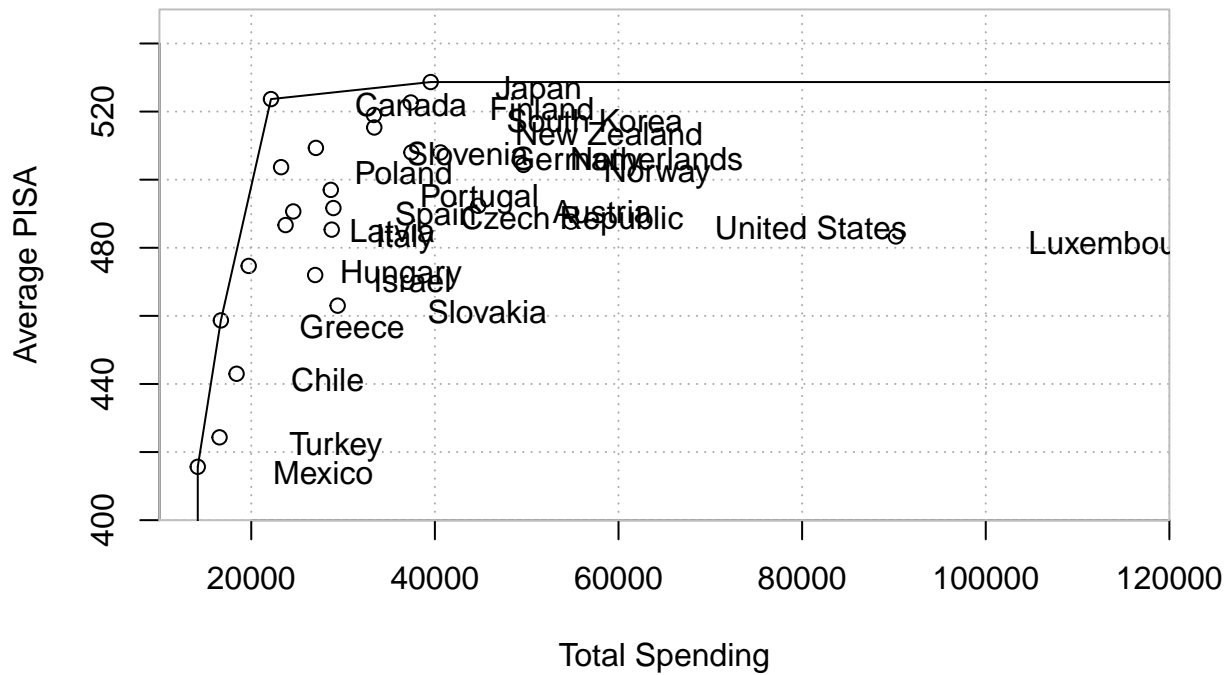
This plot is not formated for easy reading. After examining the `dea.plot()` function from Bogetoft & Otto's Benchmarking package, we were able to emulate the graphic in `ggplot`.

```
# warning, the geom_segment calls are not reproducible and must be hand specified
ggplot(DEAdata, aes(x = Total.Spending, y  = AveragePISA)) +
  geom_point() +
  geom_text(aes(label = Countries), size = 3,
            nudge_x = 2, nudge_y = -1.5, color = "gray40") +
  geom_segment(aes(x = 91000, y = 529, xend = 39541, yend = 528.67), color = "dodgerblue") +
  geom_segment(aes(x = 39541, y = 528.67, xend = 22149, yend = 523.67), color = "dodgerblue") +
  geom_segment(aes(x = 22149, y = 523.67, xend = 16691, yend = 458.67), color = "dodgerblue") +
  geom_segment(aes(x = 16691, y = 458.67, xend = 14173, yend = 415.67), color = "dodgerblue") +
  geom_segment(aes(x = 14173, y = 415.67, xend = 14173, yend = 400), color = "dodgerblue") +
  theme_minimal() +
  ggtitle("DEA plot of 2015 OECD Countries",
          subtitle = "Average Education Scores by Total Spending")
```

DEA plot of 2015 OECD Countries
Average Education Scores by Total Spending
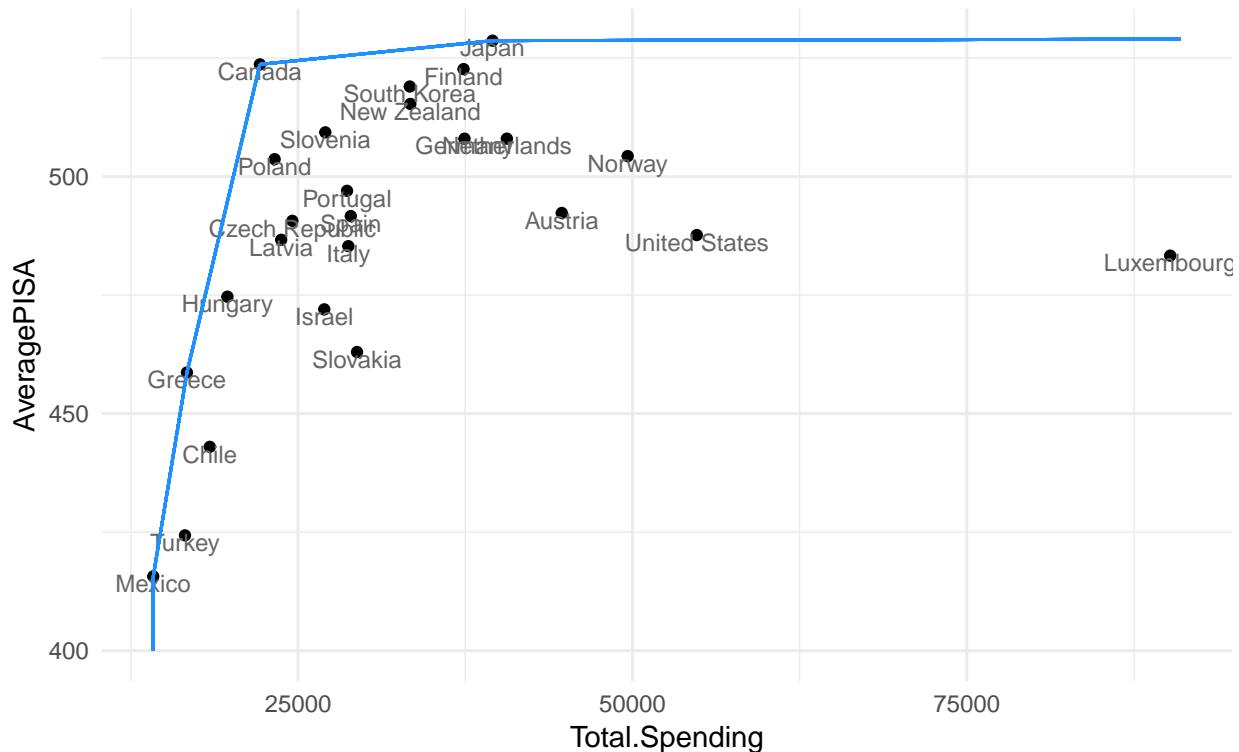

```
# ggsave("single output DEA.png") # uncomment this to update the png file for the presentation
```

Part B: Multiple Inputs and Outputs

We're also interested in how many hours each country's teachers spend teaching per year as another input, and the tertiary graduation rate as another output. Additionally, we are also using the different science, math, and reading PISA scores that made up the averagePISA input from the single input single output model. The input output diagram for this model is included in the Appendix.

Here is a model with two inputs, total educational spending per student and average teacher hours per year, and four outputs: SciencePISA, ReadingPISA, MathematicsPISA, and tertiary graduation rate.

```
# multiple input and multiple output
x <- DEAdata %>% select(Total.Spending, Teaching.Hours) ## input

y <- DEAdata %>% select(SciencePISA, ReadingPISA, MathematicsPISA, Tertiary.Graduation.Rate) ## output

row.names(x) <- DEAdata$Countries # input labels
row.names(y) <- DEAdata$Countries # output labels


resmult <- DeaMultiplierModel(x, y, rts = "vrs", orientation = "output")

#pander(resmult$Efficiency)
#pander(resmult$uy)
```
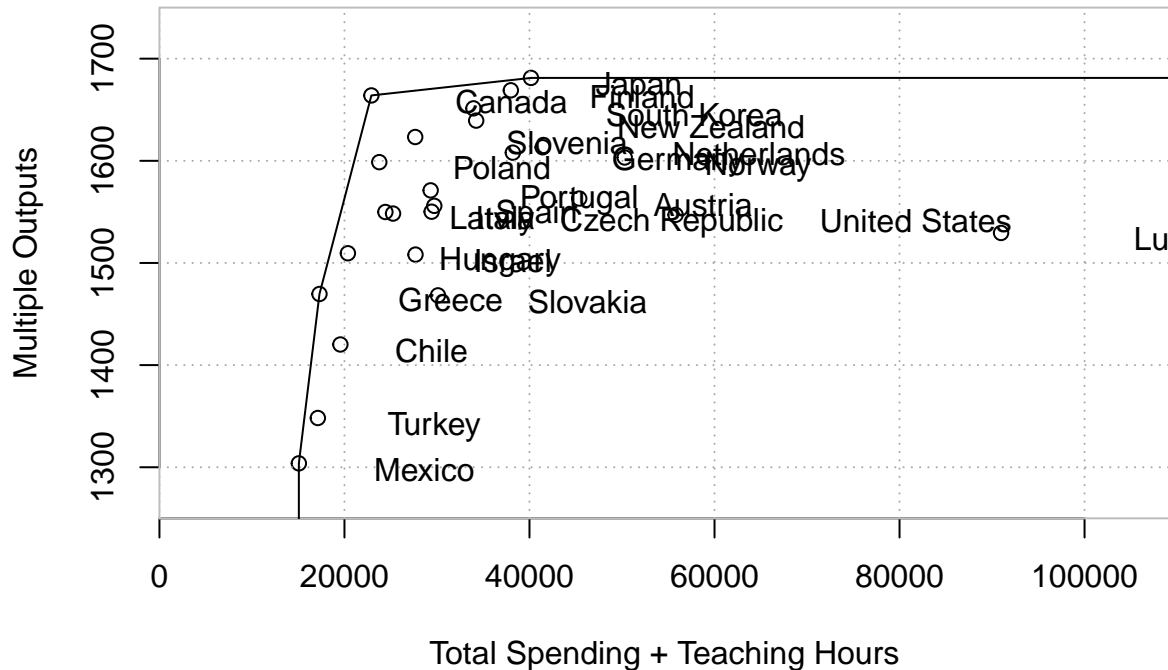
```
dea.plot(x = x, y = y, txt = dimnames(x)[[1]], GRID = T,
         xlab = "Total Spending + Teaching Hours",
         ylab = "Multiple Outputs",
         ylim = c(1250,1750), main = "2015 OECD data DEA plot")
```

2015 OECD data DEA plot


```
# warning, the geom_segment calls are not reproducible and must be hand specified
DEAdata$multIn <- DEAdata$Total.Spending + DEAdata$Teaching.Hours

DEAdata$multOut <- DEAdata$SciencePISA + DEAdata$ReadingPISA + DEAdata$MathematicsPISA +
  DEAdata$Tertiary.Graduation.Rate


ggplot(DEAdata, aes(x = multIn, y  = multOut)) +
  geom_point() +
  geom_text(aes(label = Countries), size = 3,
            nudge_x = 2, nudge_y = -2, color = "gray40") +
  geom_segment(aes(x = 91000, y = 1681, xend = 40162, yend = 1681), color = "dodgerblue") +
  geom_segment(aes(x = 40162, y = 1681, xend = 22910, yend = 1664), color = "dodgerblue") +
  geom_segment(aes(x = 22910, y = 1664, xend = 17298, yend = 1470), color = "dodgerblue") +
  geom_segment(aes(x = 17298, y = 1470, xend = 15071, yend = 1304), color = "dodgerblue") +
  geom_segment(aes(x = 15071, y = 1304, xend = 15071, yend = 1250), color = "dodgerblue") +
  theme_minimal() +
  labs(xlab = "Total Spending + Teaching Hours",
       ylab = "SciencePISA + ReadingPISA + MathematicsPISA + Tertiary Graduation Rate") +
  ggtitle("DEA plot of 2015 OECD Countries",
          subtitle = "Multiple Input and Multiple Output")
```

DEA plot of 2015 OECD Countries
Multiple Input and Multiple Output
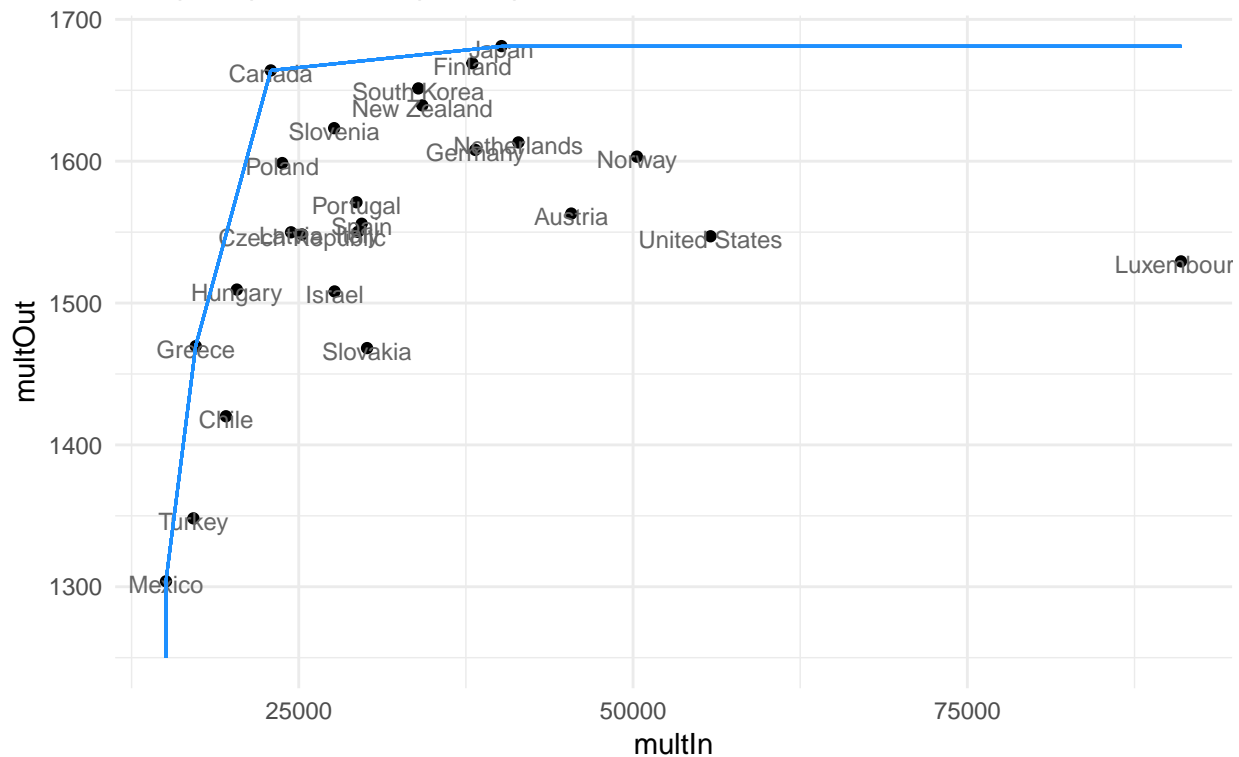

```r
ggsave("multiple output DEA.png") # uncomment this to update the png file for the presentation
```

```
## Saving 6.5 x 4.5 in image
```

Results and discussion

Single input - output results

The average efficiency for the Total Spending yielding Average PISA score is 0.952.

The table of efficiency and lambda scores indicates the following countries are output-oriented efficient units:
> Canada, Greece, Japan, Mexico

```r
df <- cbind(ressingle$Efficiency, ressingle$Lambda)

tempdf <- df[, colSums(df) != 0]



kable(tempdf, "latex", caption = "Results of Single input-Single output DEA
        displaying efficiency scores and positive lambda values",
      booktabs = T) %>%
kable_styling(latex_options = "striped", "repeat_header")
```

```r
# code for kableExtra package from Zhu(2019)
```

Mexico, Greece, Canada, Japan are 100% efficient. Slovakia is the least efficient. Mexico has the highest

Table 3: Results of Single input-Single output DEA displaying efficiency scores and positive lambda values

	Eff	Canada	Greece	Japan	Mexico
Austria	0.931	0.000	0.000	1.000	0.000
Canada	1.000	1.000	0.000	0.000	0.000
Chile	0.925	0.313	0.687	0.000	0.000
Czech Republic	0.936	0.861	0.000	0.139	0.000
Finland	0.990	0.124	0.000	0.876	0.000
Germany	0.962	0.120	0.000	0.880	0.000
Greece	1.000	0.000	1.000	0.000	0.000
Hungary	0.959	0.555	0.445	0.000	0.000
Israel	0.899	0.723	0.000	0.277	0.000
Italy	0.923	0.620	0.000	0.380	0.000
Japan	1.000	0.000	0.000	1.000	0.000
South Korea	0.985	0.356	0.000	0.644	0.000
Latvia	0.929	0.909	0.000	0.091	0.000
Luxembourg	0.914	0.000	0.000	1.000	0.000
Mexico	1.000	0.000	0.000	0.000	1.000
Netherlands	0.961	0.000	0.000	1.000	0.000
New Zealand	0.978	0.353	0.000	0.647	0.000
Norway	0.954	0.000	0.000	1.000	0.000
Poland	0.961	0.937	0.000	0.063	0.000
Portugal	0.946	0.625	0.000	0.375	0.000
Slovakia	0.881	0.582	0.000	0.418	0.000
Slovenia	0.970	0.719	0.000	0.281	0.000
Spain	0.935	0.609	0.000	0.391	0.000
Turkey	0.930	0.000	0.942	0.000	0.058
United States	0.922	0.000	0.000	1.000	0.000

output weightage followed by Turkey. The target of performance for Turkey (w.r.t Mexico) is Mexico scaled down by a factor of 0.0576.

The lambda values of Japan against other countries ranges from 0 to 1. The countries with 0 lambda values (Canada, Chile, Greece, Hungary, Mexico, Turkey lie on or close to the efficiency frotier) cannot be compared with Japan. Austria, Luxembourg, Netherlands, Norway and U.S. have lambda values of 1 meaning the output of these countries cannot be scaled higher than Japan for the same level of input to achieve 100% efficiency. The input values (Spending per student) of these countries are already much higher compared to Japan.

Multiple input - output results

The average efficiency for the Total Spending yielding Average PISA score is 0.972.

The table of efficiency and lambda scores indicates the following countries are output-oriented efficient units:
> Canada, Finland, Greece, Japan, South Korea, Poland, Slovenia

```
df <- cbind(resmult$Efficiency, resmult$Lambda)

tempdf <- df[, colSums(df) != 0]

kable(tempdf, "latex", caption = "Results of Multiple Input Multiple Output DEA
```

Table 4: Results of Multiple Input Multiple Output DEA displaying efficiency scores and positive lambda values

	Eff	Canada	Finland	Greece	Japan	South Korea	Mexico	Poland	Slovenia	Turkey
Austria	0.938	0.119	0.000	0.000	0.881	0.000	0	0.000	0.000	0
Canada	1.000	1.000	0.000	0.000	0.000	0.000	0	0.000	0.000	0
Chile	0.969	0.041	0.072	0.887	0.000	0.000	0	0.000	0.000	0
Czech Republic	0.955	0.602	0.000	0.000	0.000	0.197	0	0.202	0.000	0
Finland	1.000	0.000	1.000	0.000	0.000	0.000	0	0.000	0.000	0
Germany	0.972	0.707	0.000	0.000	0.293	0.000	0	0.000	0.000	0
Greece	1.000	0.000	0.000	1.000	0.000	0.000	0	0.000	0.000	0
Hungary	0.984	0.374	0.000	0.475	0.000	0.000	0	0.151	0.000	0
Israel	0.953	0.229	0.436	0.335	0.000	0.000	0	0.000	0.000	0
Italy	0.971	0.354	0.477	0.142	0.000	0.000	0	0.000	0.027	0
Japan	1.000	0.000	0.000	0.000	1.000	0.000	0	0.000	0.000	0
South Korea	1.000	0.000	0.000	0.000	0.000	1.000	0	0.000	0.000	0
Latvia	0.954	0.503	0.054	0.135	0.000	0.000	0	0.000	0.308	0
Luxembourg	0.924	0.392	0.000	0.000	0.608	0.000	0	0.000	0.000	0
Mexico	1.000	0.000	0.000	0.000	0.000	0.000	1	0.000	0.000	0
Netherlands	0.970	0.249	0.000	0.000	0.751	0.000	0	0.000	0.000	0
New Zealand	0.996	0.353	0.000	0.000	0.647	0.000	0	0.000	0.000	0
Norway	0.977	0.224	0.685	0.000	0.092	0.000	0	0.000	0.000	0
Poland	1.000	0.000	0.000	0.000	0.000	0.000	0	1.000	0.000	0
Portugal	0.956	0.399	0.246	0.000	0.114	0.051	0	0.190	0.000	0
Slovakia	0.909	0.565	0.000	0.000	0.386	0.049	0	0.000	0.000	0
Slovenia	1.000	0.000	0.000	0.000	0.000	0.000	0	0.000	1.000	0
Spain	0.941	0.986	0.000	0.000	0.014	0.000	0	0.000	0.000	0
Turkey	1.000	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000	1
United States	0.943	1.000	0.000	0.000	0.000	0.000	0	0.000	0.000	0

```
    displaying efficiency scores and positive lambda values", booktabs = T) %>%
kable_styling(latex_options = c("striped", "scale_down", "repeat_header"))

# code for kableExtra package from Zhu(2019)
```

It seems that with these changed inputs, Canada, Mexico and Japan remain efficient but several other countries, including Turkey, Slovenia, and Poland have become fully efficient, implying that they do a good job utilizing their teacher's hours efficiently. Greece has the highest weight applied to its graduation rate, implying that they do relatively well on this measure.

References

[1] Anderson, T. R. (2019) Operations Research Using R

[2] Anderson, T. R. (2019) Data Envelopment Analysis Using R

[3] Bogetoft, P., Otto, L. (2011) Benchmarking with DEA, SFS, and R. Springer.

[4] Gavurova, B., Kocisova, K., Belas, L., & Krajcik, V. (2017). Relative efficiency of government expenditure on secondary education. Journal of International Studies, 10(2), 329-343. doi:10.14254/2071-8330.2017/10-2/23

[5] Zhu, H. (2019) Create Awesome LaTeX Table with knitr::kable and kableExtra. https://haozhu233.github.io/kableExtra/awesome_table_in_pdf.pdf

[6] John D. Lamb, Kai-Hong Tee, Data envelopment analysis models of investment funds, European Journal of Operational Research, Volume 216, Issue 3, 2012, Pages 687-696

[7]"Data Envelopment Analysis Explained." Service Productivity Management: Improving Service Performance Using Data Envelopment Analysis (DEA), by H. David. Sherman and Joe Zhu, Springer Science, 2006.
[8] data envelopment analysis, YOng bae Ji and Choonjoo Lee

[9]José L. Ruiz, José V. Segura, Inmaculada Sirvent, Benchmarking and target setting with expert preferences: An application to the evaluation of educational performance of Spanish universities, European Journal of Operational Research, Volume 242, Issue

Data Sources

Organization for Economic Development https://data.oecd.org/ UNESCO (add URL here) Teacher hours data source https://data.oecd.org/eduresource/teaching-hours.htm

Appendix

Data Summary

```
glimpse(data)
```

```
## Observations: 40
## Variables: 22
## $ Countries                  <chr> "Australia", "Austria", "Belgium...
## $ GDP                        <dbl> 1349.0, 382.1, 455.0, 1560.0, 24...
## $ SciencePISA                <int> 510, 495, 502, 528, 447, 493, 50...
## $ ReadingPISA                <int> 503, 485, 499, 527, 459, 487, 50...
## $ MathematicsPISA            <int> 494, 497, 507, 516, 423, 492, 51...
## $ AveragePISA                <dbl> 502, 492, 503, 524, 443, 491, 50...
## $ HDI                        <dbl> 0.936, 0.903, 0.913, 0.920, 0.84...
## $ Primary.education          <chr> "9,546", "11,689", "10,211", "9,...
## $ Secondary.education        <chr> "12,303", "15,477", "13,070", "1...
## $ Tertiary.education         <chr> "20,344", "17,555", "17,320", "0...
## $ Total.Spending             <dbl> 42193, 44721, 40601, 22149, 1840...
## $ Tertiary.Graduation.Rate   <dbl> NA, 86.1, 31.4, 93.0, 91.2, 76.4...
## $ Bachelors.Graduation.rates <dbl> 59.77, 25.01, 43.90, 37.59, 35.9...
## $ Masters                    <dbl> 20.53, 20.29, 26.76, 11.78, 10.1...
## $ Doctorates                 <dbl> 2.618, 1.862, 0.639, 1.559, 0.26...
## $ Spending.as.percentage.of.GDP <chr> "5.32%", "5.45%", "6.55%", "", "...
## $ Teaching.Hours            <dbl> 825, 658, NA, 761, 1157, 676, NA...
## $ X                         <lgl> NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ X.1                       <lgl> NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ X.2                       <lgl> NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ X.3                       <lgl> NA, NA, NA, NA, NA, NA, NA, NA, ...
## $ X.4                       <lgl> NA, NA, NA, NA, NA, NA, NA, NA, ...
```

Let's glance at the teaching hours and the graduation rate:

```
head(cbind(data$Countries, data$Teaching.Hours, data$Tertiary.Graduation.Rate))
```

```
##      [,1]          [,2]           [,3]
## [1,] "Australia"   "825.4093373" NA
```

```
## [2,] "Austria"        "658.2"       "86.079"
## [3,] "Belgium"        NA            "31.35"
## [4,] "Canada"         "760.623341"  "93.005"
## [5,] "Chile"          "1157.36"     "91.196"
## [6,] "Czech Republic" "676.3166667" "76.361"
```

Code for DEA Model Graphs

For some reason the DiagrammeR graphs won't render to pdf directly, only to html (then print to pdf). They can be previewed and saved as png in RStudio IDE, which is what we did for this report.

Single input single output DEA model

```
grViz("
digraph nicegraph {

  # a 'graph' statement
  graph [overlap = true, fontsize = 10]

  # several 'node' statements
  node [shape = rectangle, fixedsize = true, width = 1,
        color = darkslategray]
  Eff

  node [shape = plaintext]
  TotalSpending; AveragePISA

  # 'edge' statements
  edge [color = grey]
  TotalSpending->Eff Eff->AveragePISA
}
")
```

Multiple input multiple output model

```
grViz("
digraph nicegraph {

  # a 'graph' statement
  graph [overlap = true, fontsize = 10]

  # several 'node' statements
  node [shape = rectangle, fixedsize = true, width = 1,
        color = darkslategray]
  Eff

  node [shape = plaintext]
  TotalSpending; TeachingHours;
  SciencePISA; ReadingPISA; MathPISA; TertGradRate

  # 'edge' statements
  edge [color = grey]
  TotalSpending->Eff TeachingHours->Eff Eff->TertGradRate
```

TotalSpending

Eff

AveragePISA
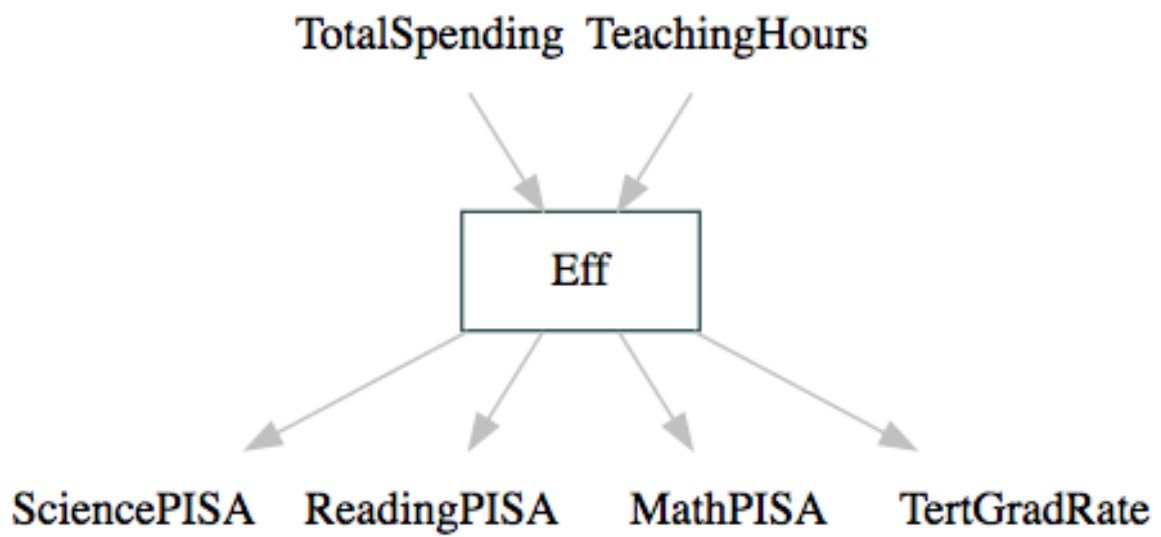
Figure 1: Single Input Single Output

Figure 2: Multiple Input Multiple Output

```
  Eff->SciencePISA Eff->ReadingPISA Eff->MathPISA
}
")
```

Deleted sections

Part B:

Run a DEA with a single input, but multiple outputs:

Alternative Intro Materials

Education has long been identified as a catalyst for human development. The rise of every civilization has coincided with a shift in attitude towards attaining and valuing knowledge, and the peak of each civilization usually aligns with peaks in their attained knowledge. For this reason, most historical societies tried to place an emphasis in the process of individual citizens attaining the available knowledge to maximize an individual's contributions to society. This process, however, was informal and was only available to a select few.

After the renaissance period, the approach to education process was far or less standarized. Societies realized the need to produce educated individuals that can tackle new problems faced by society and governments throughout the world, started to shape their educational curriculum's based on the country's need. Today, most countries have curricula that are directly based on this but adapted to fit their national goals and cultural values. The focus and the mission of each country's educational administration directly reflects their future aspirations and social, political, and economical goals.

For example, the United States Department of Education's mission statement includes "to promote student achievement and preparation for global competitiveness." In Germany, the education department is grouped with research and controlled under the department of Education and Research. Their mission includes "promote education and research for they are the foundations on which we will build our future in a changing world". In Mexico, the department's goal is "to create general conditions that permit the assurance of quality education to all Mexicans". These differing mission objectives, though all noble, point out the subtle differences each society attaches to the education process.

The subtle differences can also translate to funding policies. Currently, budgets for Education Departments around the world constitues a good percentage of each countries overall GDP. For countries within the "Organization for Economic Development" (OECD), current budgets stand between 4-8 percent of their entire GDP. Globally, the current spending on education alone is estimated at ~5 trillion U.S dollars with this figure estimated to reach ~10 trillion dollars in the year 2030. This is a huge investment without any short term returns. For politicians whose lifespan in the political arena is limited (*Democratic Soceities*), such spending would need justification and buy in.

Additional extra code, ok to delete if not needed for final Thursday deadline

Some quick analysis of the results:

Note that we're missing quite a few elements of data for tertiary graduation rate.