# Productivity and Efficiency Analysis

## 2) Data envelopment analysis (DEA)

*a) Application of DEA*

**Timo Kuosmanen**

Aalto University School of Business

https://people.aalto.fi/timo.kuosmanen

# Cost frontier model (2012-2015)

$$\ln x = \ln C(y_1, y_2, y_3) + \delta z + u + v$$

- $x$ is the total cost
- $C$ is the frontier cost function
- $y_1$ is the energy transmission (GWh of 0.4 kV equivalents)
- $y_2$ is the total length of the network (km)
- $y_3$ is the number of customers connected to the network
- $z$ is the proportion of underground cables (%)
- $u$ is an asymmetric inefficiency term
- $v$ is a random noise term

# Simplified DEA model

$$\ln x = \ln C(y_1, y_2, y_3) + u$$

- $x$ is the total cost
- $C$ is the frontier cost function
- $y_1$ is the energy transmission (GWh of 0.4 kV equivalents)
- $y_2$ is the total length of the network (km)
- $y_3$ is the number of customers connected to the network
- ~~$z$ is the proportion of underground cables (%)~~
- $u$ is an asymmetric inefficiency term
- ~~$v$ is a random noise term~~

# DEA model specification

To apply DEA, user must first make the following specifications

- Input variables **X**
- Output variables **Y**
- Returns to scale: Constant, variable (non-increasing, non-decreasing)
- Orientation: input orientation (saving mode), output orientation (expansion mode)

# DEA results

There are three types of information that we can get from DEA

- Efficiency scores of each firm
- Multiplier weights $\boldsymbol{\gamma}$ (output weights)
- Intensity weights $\boldsymbol{\lambda}$ (benchmark firms)

# Computing DEA in R

The following examples are based on the R tutorial prepared by Sheng Dai, available on the course website.

Further reading:

- [1] Bogetoft P, Otto L. Benchmarking with DEA, SFA, and R. Springer; 2010.
- [2] Behr A. Production and efficiency analysis with R. Springer; 2015.

# DEA efficiency scores: CRS, input orientation

Example: the first 8 firms:

```
##          eff eff.d
##  [1,] 0.756 0.756
##  [2,] 0.805 0.805
##  [3,] 0.645 0.645
##  [4,] 0.807 0.807
##  [5,] 0.693 0.693
##  [6,] 0.907 0.907
##  [7,] 0.815 0.815
##  [8,] 0.690 0.690
```

Average Efficiency: 80.2%

Standard Deviation: 11.9%

# Multiplier weights γ (output weights)

Example: the first 8 firms, the highest weight indicated by red rectangle:

| v1 | v2 | v3 |
|---|---|---|
| 3.52e-03 | 5.10e-04 | 9.01e-06 |
| 0.00e+00 | 4.80e-04 | 5.57e-05 |
| 0.00e+00 | 4.66e-04 | 5.41e-05 |
| 2.79e-04 | 4.92e-05 | 0.00e+00 |
| 1.39e-03 | 9.41e-04 | 0.00e+00 |
| 1.03e-03 | 1.39e-04 | 2.05e-05 |
| 3.95e-03 | 6.99e-04 | 0.00e+00 |
| 1.61e-03 | 2.34e-04 | 4.13e-06 |

# Intensity weights λ (benchmarks)

Example: the first 8 firms. There are 8 firms that are 100% efficient (#22, #28, #32, #37, #46, #56, #70, #73). These 8 firms serve as benchmarks for all other units.

| L22 | L28 | L32 | L37 | L46 | L56 | L70 | L73 |
|---|---|---|---|---|---|---|---|
| 0.01987 | 0.000000 | 0.00995 | 0.00000 | 0.0000 | 0.0000 | 0.20775 | 0.00000 |
| 0.21307 | 0.000000 | 0.00000 | 0.19495 | 0.0000 | 0.0000 | 0.00000 | 0.00000 |
| 0.32064 | 0.000000 | 0.00000 | 0.08602 | 0.0000 | 0.0000 | 0.00000 | 0.00000 |
| 0.00000 | 0.000000 | 0.00000 | 2.40268 | 0.0000 | 0.0000 | 0.86121 | 0.00000 |
| 0.00000 | 0.026182 | 0.00000 | 0.00708 | 0.0000 | 0.0000 | 0.00000 | 0.00000 |
| 0.26732 | 0.000000 | 0.00000 | 0.00000 | 0.0000 | 0.1283 | 0.00000 | 0.06900 |
| 0.00000 | 0.000000 | 0.00000 | 0.22793 | 0.0000 | 0.0000 | 0.01590 | 0.00000 |
| 0.28274 | 0.000000 | 0.03214 | 0.00000 | 0.0000 | 0.0000 | 0.30727 | 0.00000 |

# DEA efficiency scores: CRS vs VRS, input orientation

Example: the first 8 firms:

**CRS**

```
##          eff eff.d
##  [1,] 0.756 0.756
##  [2,] 0.805 0.805
##  [3,] 0.645 0.645
##  [4,] 0.807 0.807
##  [5,] 0.693 0.693
##  [6,] 0.907 0.907
##  [7,] 0.815 0.815
##  [8,] 0.690 0.690
```

**VRS**

```
##          eff eff.d
##  [1,] 0.767 0.767
##  [2,] 0.815 0.815
##  [3,] 0.655 0.655
##  [4,] 0.834 0.834
##  [5,] 0.730 0.730
##  [6,] 0.912 0.912
##  [7,] 0.847 0.847
##  [8,] 0.692 0.692
```

Note: VRS efficiency score is always greater than or equal to CRS efficiency

# Intensity weights λ (benchmarks) in VRS

Example: the first 8 firms. In VRS case, there are 11 firms that are 100% efficient (#12, #15, #22, #28, #32, #37, #46, #56, #61, #70, #73). These 11 firms serve as benchmarks for all other units.

Note: intensity weights λ must sum to 1 in VRS.

| L12 | L15 | L22 | L28 | L32 | L37 | L46 | L56 | L61 | L70 | L73 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0000 | 0.0000 | 0.00000 | 0.00000 | 0.000000 | 0.0169 | 0.0000 | 0.000000 | 0.8114 | 0.17170 | 0.00000 |
| 0.0000 | 0.0000 | 0.15485 | 0.00000 | 0.000000 | 0.1888 | 0.0000 | 0.000000 | 0.6563 | 0.00000 | 0.00000 |
| 0.0000 | 0.0000 | 0.26228 | 0.00000 | 0.000000 | 0.0799 | 0.0000 | 0.000000 | 0.6578 | 0.00000 | 0.00000 |
| 0.0000 | 0.0000 | 0.00000 | 0.41232 | 0.106158 | 0.0000 | 0.0000 | 0.000000 | 0.0000 | 0.43117 | 0.05036 |
| 0.0000 | 0.0000 | 0.00000 | 0.00000 | 0.000000 | 0.0233 | 0.5079 | 0.000000 | 0.4688 | 0.00000 | 0.00000 |
| 0.0000 | 0.0000 | 0.15870 | 0.00000 | 0.000000 | 0.0000 | 0.0000 | 0.146362 | 0.6320 | 0.00000 | 0.06289 |
| 0.0000 | 0.0000 | 0.00000 | 0.00000 | 0.000000 | 0.1418 | 0.2889 | 0.000000 | 0.5693 | 0.00000 | 0.00000 |
| 0.0000 | 0.0000 | 0.24774 | 0.00000 | 0.030439 | 0.0000 | 0.0000 | 0.000000 | 0.4186 | 0.30318 | 0.00000 |

# Some basic properties of DEA

- VRS efficiency is always greater than or equal to the CRS efficiency
- 100% efficient units in CRS are also 100% efficient in VRS.
- Scale efficiency = CRS efficiency / VRS efficiency

- In CRS, input-oriented efficiency score

  = 1/ output-oriented efficiency
- In VRS, input and output orientations can yield different results

# Advantages of DEA

- Data-driven approach: does not depend on any arbitrary parametric assumptions
- Axiomatic foundation: DEA enforces the right curvature of the frontier
  - Monotonicity and convexity of output isoquants are important properties in incentive regulation

- *Simple to use: advantage or disadvantage?*

# Disadvantages of DEA

- Frontier is determined by just a handful of efficient units: “a great waste of information” (Farrell, 1957)
- Any omitted factors, unobserved heterogeneity, or data errors in the evaluated unit or in the benchmarks will be directly attributed as ”inefficiency”: DEA is sensitive to noise
- If the number of input and outputs increases, a large proportion of units will appear as “efficient”: *the curse of dimensionality*

# Next lesson

2b) DEA formulations