Inroduction to Special Topics in Business Economics

"Technical Efficiency (TE) in R"
Using Data Envelopment Analysis (DEA) under different technology assumptions

PhD Candidate: Eirini Stergiou
e.stergiou@upnet.gr

School of Business Administration
Department of Economics
University of Patras, Greece

November 10, 2018

Structure

1. Efficiency measure
   - Introduction
   - Benchmarking package
   - One Input - One Output
   - Two Inputs - One Output
   - Example with a dataset

- **Input** orientation:"By how much can input quantities be proportionally reduced without changing the output quantities produced?"
- **Output** orientation:"By how much can output quantities be proportionally increased without changing the input quantities used?"
- CRS - VRS assumption
- One advantage of the Farrell input and output orientated **radial** TE measures is that they are *units invariant*
- DEA is the non-parametric mathematical programming approach to frontier estimation
  - Charnes, Cooper and Rhodes (1978) proposed a model which had an input orientation and assumed CRS
  - Banker, Charnes and Cooper (1984) proposed a VRS model
  - Second stage of linear programming by maximizing the sum of **slacks** (e.g. Ali and Seiford (1993))
    However some problems occur (furthest efficient point,not unit invariant)

We use the **Benchmarking** package

- It contains an extensive number of methods for parametric and nonparametric efficiency analysis
- It offers a variety of DEA methods and it is easy to use plotting facilities

- x: Labor
- y: Sales
- CRS - calculate the ratio of output to input
- DMU B has the highest ratio
- Dividing all ratios by the maximal ratio $\Longrightarrow$ efficiency scores

	A	B	C	D	E
Input x	1.000	2.000	3.000	4.000	5.000
Output y	1.000	3.000	2.000	5.000	4.000
y/x	1.000	1.500	0.667	1.250	0.800
Efficiency	0.667	1.000	0.444	0.833	0.533

We can benchmark all DMUs indexed by i relative to the efficient DMU B

$$0 \leq \frac{\text{Sales per employee of DMU i}}{\text{Sales per employee of B}} \leq 1 \qquad (1)$$

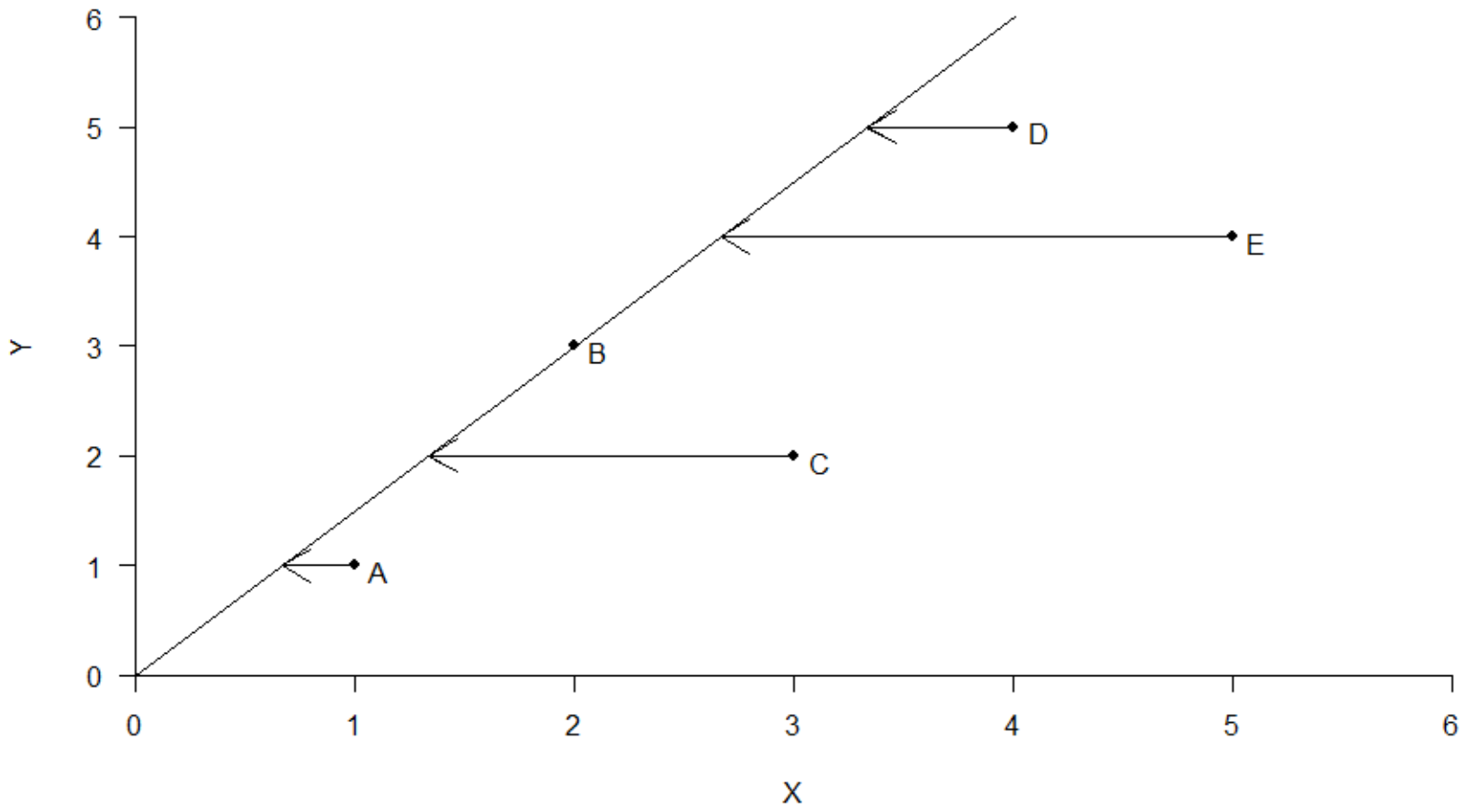

Figure 1: One input - one output case

Coding

- install.packages("Benchmarking")
- library(Benchmarking)
- emp <- matrix(1:5)
- sales <- matrix(c(1,3,2,5,4))
- nam <- LETTERS[1:5]
- maxc<-cmax(sales/emp)
- eff <- sales/emp / max
- tab <- t(round(cbind(emp,sales,round(sales/emp,3),eff),3))
- colnames(tab) <- nam
- rownames(tab) <- c("input x","output y","y/x","efficiency")
- View(tab)
- **dea.plot(emp,sales,RTS="crs",ORIENTATION="in-out",pch=19,cex=0.8,txt=LETTERS[1:length(emp)],las=1)**

- x_1: Employees
- x_2: Floor area
- y: Sales
- CRS
- Isoquant curve

	A	B	C	D	E
Input x_1	4.000	7.000	8.000	5.000	2.000
Input x_2	3.000	2.000	1.000	5.000	5.000
Output y	1.000	1.000	1.000	1.000	1.000
Efficiency	1.000	0.909	1.000	0.700	1.000

The three efficient units A, C, and E serve as the benchmark for the non-efficient DMUs.

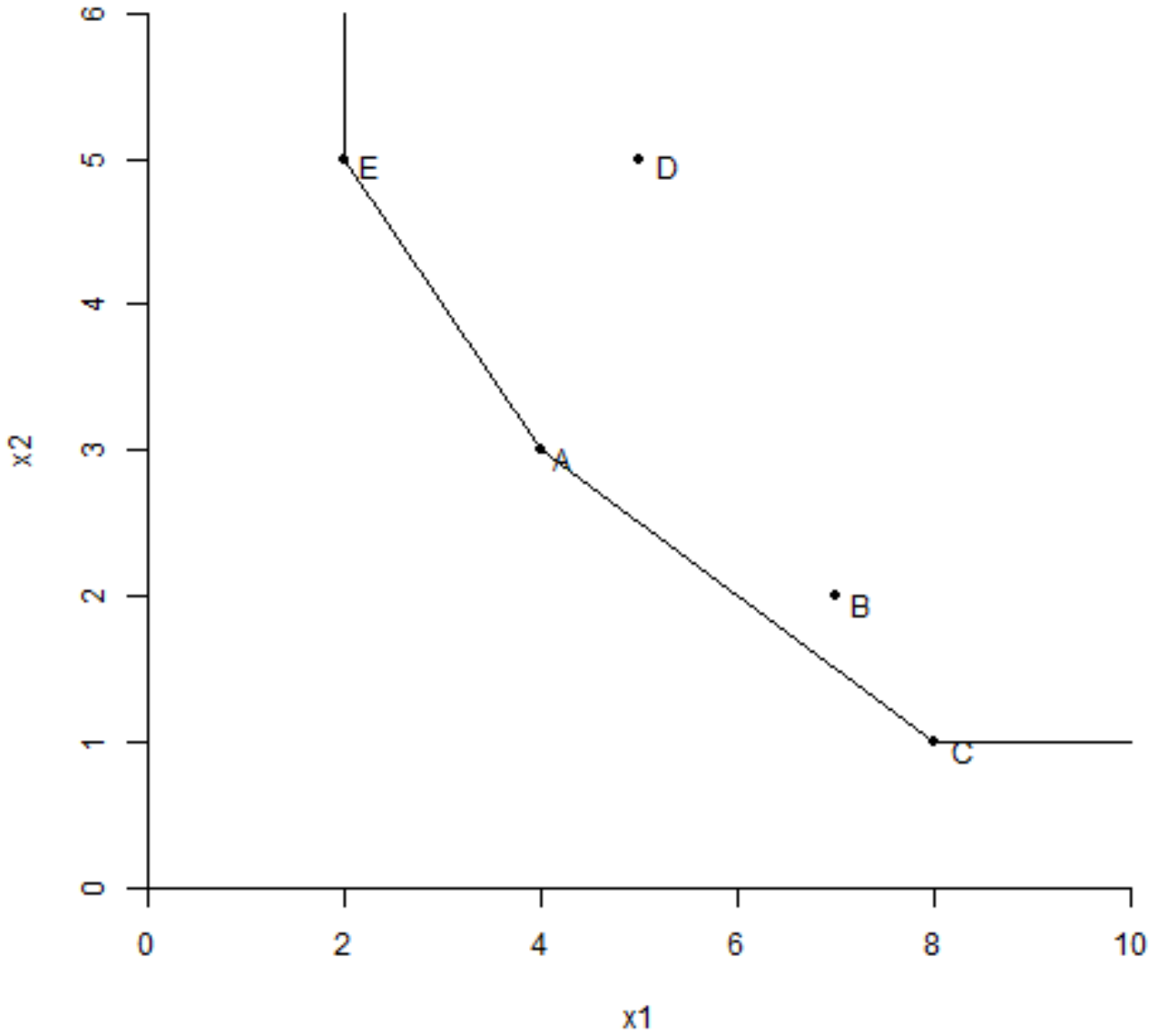

Figure 2: Two inputs - one output case

Coding

- x1 <- c(4,7,8,5,2)
- x2 <- c(3,2,1,5,5)
- y <- rep(1,5)
- X <- cbind(x1,x2)
- Y <- matrix(y)
- **e <- dea(X,Y,RTS="crs",ORIENTATION="in",SLACK=T)**
- eff(e)
- **plot <- dea.plot.isoquant(x1,x2,txt=LETTERS[1:5], xlim=c(0,10),pch=19,cex=0.8)**

Dataset charnes1981

A data frame with 70 school sites

- firm: school site number
- x1: education level of the mother
- x2: highest occupation of a family member
- x3: parental visits to school
- x4: time spent with children in school-related topics
- x5: the number of teachers at the site
- y1: reading score
- y2: math score
- y3: self - esteem score
- pft: =1 if in program (program follow through) and =0 if not in program
- name: Site name

Coding

- *Reading a file*
  - charnes1981 <- read.csv("C:/Users/Irene/Documents/R/win-library/3.4/Benchmarking/data/charnes1981.csv/charnes1981.csv", header=TRUE, sep=";")
- x <- with(charnes1981, cbind(x1,x2,x3,x4,x5))
- y <- with(charnes1981, cbind(y1,y2,y3))
- **Phase one: Farrell input efficiency - vrs technology**
  - e <-dea(x, y, RTS="vrs", ORIENTATION="in")
  - peers(e)
  - lambda(e)
  - summary(e)
  - excess(e,x)
- **Phase two: Calculate slacks (maximize sum of slacks)**
  - el <- dea(x,y,SLACK=TRUE)
- *Creating a data frame*
  - total <- data.frame(eeff,elslack,elsx,elsy)
- *Saving the efficiency results*
  - write.csv(total, file = "C:/Users/Irene/Desktop/efficiency.csv")

References

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. Management science, 30(9), 1078-1092.
Charnes, A., Cooper, W. W., & Rhodes, E. (1981) Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through. Management Science, 27(6), 668-697.
Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. European journal of operational research, 2(6), 429-444.
Peter Bogetoft and Lars Otto. Benchmarking with DEA, SFA, and R, Springer 2011. Sect. 5.2 page 115
P Andersen and NC Petersen. "A procedure for ranking efficient units in data envelopment analysis", Management Science 1993 39(10):1261-1264
`https://cran.r-project.org/web/packages/Benchmarking/Benchmarking.pdf`