



UNIVERSITY OF COPENHAGEN

Introduction to SFA and DEA using R

Arne Henningsen & Tomasz G. Czekaj
Department of Food and Resource Economics

15 June 2015

Programme

- 9:00 – 10:00 general introduction to R
- 10:00 – 10:30 SFA using R
- 10:30 – 11:00 coffee break
- 11:00 – 11:45 SFA using R (cont.)
- 11:45 – 12:30 SFA and DEA using R



Some Information on R

- 20+ years old
- developed by leading statisticians
- free & open-source
- all major platforms and OSs
- *lingua franca* in statistics
- interactive command-line interface
- minimalistic GUI, R GUIs, RStudio IDE
- script files
- functional & object-oriented
- easily extensible
- more than 6,500 add-on packages
- useRs become developErs



Manuals, Tutorials, ...

- official manuals, e.g.
 - Introduction to R
 - R Data Import/Export
 - Writing R Extensions
- “contributed documentation”
- other internet sources (texts, videos, ...)
- books
 - A Beginner’s Guide to R (Zuur, Ieno & Meesters, 2009)
 - R Cookbook (Teetor, 2011)
 - Applied Econometrics with R (Kleiber & Zeileis, 2008)
 - ...
- MOOCs, e.g. at Coursera
- ...



General Introduction to R

- CRAN: R packages
- task views:
 - econometrics
 - time series analysis
 - optimisation and mathematical programming
 - statistics for social sciences
 - analysis of spatial data



General Introduction to R

- start RStudio IDE
- different panels of the RStudio window
- brief introduction to R language



Exercises:

- Exercise 1: Testing Linear Models
- Exercise 2: Panel Data



Exercise 1: Testing Linear Models

- 1 load the data set `riceProdPhil` (package “`frontier`”)
- 2 estimate a Cobb-Douglas production function with two inputs: land and labour
- 3 estimate a Cobb-Douglas production function with four inputs: land, labour, fertilizer, and other inputs
- 4 compare the two models: t -tests, Wald test, LR test (using `waldtest()` and `lrtest()` from package “`lmtest`”)



Exercise 2: Panel Data

- 1 load the data set `riceProdPhil` (package “`frontier`”)
- 2 estimate a Cobb-Douglas production function with three inputs (land, labour, fertilizer) and fixed individual effects (using `plm()` from package “`plm`”)
- 3 estimate a Cobb-Douglas production function with three inputs and random individual effects (using argument “`model`” of `plm()`)
- 4 test the consistency of the random-effects estimator with a Hausman test (using `phptest()` from package “`plm`”)



Productivity and Efficiency Analysis with R

R packages

- “**frontier**” (based on FRONTIER 4.1)
- “sfa” (simple SFA)
- “spfrontier” & “ssfa” (spatial SFA)
- “semsfa” (semiparametric SFA)
- “**Benchmarking**” & “FEAR” & “nonparaeff” (DEA)
- “rDEA” (robust DEA)
- “TFDEA” (technology forecasting using DEA)
- “frontiles” (alpha-quantile and order-m efficiency)



Exercises:

- Exercise 3: Cross-Sectional/Pooled SFA
- Exercise 4: Panel SFA
- Exercise 5: DEA



Exercise 3: Cross-Sectional/Pooled SFA

- 1 load the data set `riceProdPhil` (package “`frontier`”)
- 2 estimate a pooled Cobb-Douglas stochastic production frontier with three inputs (land, labour, fertilizer)
(using `sfa()` from package “`frontier`”)
- 3 what is the elasticity of scale?
- 4 what proportion of the variance of the residual is due to noise and what proportion is due to inefficiency?
- 5 plot a histogram of the estimated technical efficiencies
- 6 which observation (farm, year) has the lowest efficiency?
- 7 use a LR test to check if the fit of the SFA model is significantly better than the fit of an OLS model



Exercise 4: Panel SFA

- 1 load the data set `riceProdPhil` (package “`frontier`”)
- 2 estimate a Cobb-Douglas stochastic production frontier with three inputs (land, labour, fertilizer) and time-invariant technical efficiency
- 3 compare the efficiencies from this model with the efficiencies from the pooled model
- 4 estimate a Cobb-Douglas stochastic production frontier with three inputs (land, labour, fertilizer) and time-varying technical efficiency
- 5 use a LR test to check if the efficiencies vary over time
- 6 estimate a ‘true fixed effects’ Cobb-Douglas stochastic production frontier with three inputs (land, labour, fertilizer) and observation-specific technical efficiency
- 7 use a LR test to test the statistical significance of the fixed individual effects



Exercise 5: DEA

- 1 load the data set `riceProdPhil` (package “`frontier`”)
- 2 load package `Benchmarking`
- 3 solve input-oriented DEA with VRS and CRS (input and output oriented)
(using `dea()` from package “`Benchmarking`”)
- 4 calculate scale efficiencies and investigate economies of scale
- 5 calculate super-efficiency scores
(using `sdea()` from package “`Benchmarking`”)
- 6 obtain and decompose productivity changes (for simplicity only between the first two time periods in the dataset) (using arguments “`XREF`” & “`YREF`” of `dea()` to specify the “reference” technologies, e.g. for period $t=0$, $t=1$ etc.)

