Efficiency Analysis in R using Parametric, Semiparametric, and Nonparametric Methods

Arne Henningsen[1,2,*], **Subal Kumbhakar**[3]

1. Department of Agricultural Economics, University of Kiel (Germany)
2. Institute of Food and Resource Economics, University of Copenhagen (Denmark)
3. Department of Economics, State University of New York at Binghamton (USA)

* Contact author: arne.henningsen@gmail.com

Keywords: Efficiency, Productivity, Parametric, Semiparametric, Nonparametric

Efficiency and productivity analysis is a major field in applied production economics. It is generally dominated by two methods: the parametric Stochastic Frontier Analysis (SFA) and the nonparametric and deterministic Data Envelopment Analysis (DEA). The SFA can be done in R with the **frontier** package [1] and the DEA might be done with the **FEAR**[1] package [4]. The SFA approach contains a stochastic error term and hence, is suitable even if there is some "noise" in the data. However, this parametric approach requires the specification of an explicit functional form, although the functional form cannot be derived from theory. Selecting a wrong functional form may lead to severely biased estimation results. If the data set includes production units with rather different technologies, even flexible functional forms cannot model their production technologies adequately and hence, the parametric SFA is inappropriate. In contrast, the nonparametric and deterministic DEA approach does not require the specification of a functional form, but it does not include a stochastic component. Hence, the DEA is not suitable in case of "noisy" data.

In many real world applications, the data are noisy *and* production units have rather different technologies (in parametric sense) so that a stochastic and nonparametric approach is required and neither the SFA nor the DEA is appropriate. In cases like this, a semiparametric SFA [2] is appropriate, because it allows for statistical "noise" and does not require the specification of a functional form for production technologies. In a first step, a nonparametric production function is estimated and in a second step the residuals of the first step are used to estimate inefficiencies. Although in many empirical applications this approach seems to be more appropriate than the SFA and DEA, it has not been used much in applied studies, probably because of nonavailability of user-friendly software. However, several soft-ware packages for nonparametric econometrics have become available in recent years. For instance, the powerful and feature-rich **np** package [3] can be used in the first step to estimate the nonparametric production function and the **frontier** package [1] can be used in the second step to estimate the technical efficiencies.

We will demonstrate how the three approaches (SFA, DEA, and semiparametric SFA) can be used for applied efficiency analysis in **R** and we compare the results obtained from all three approaches.

References

[1] Tim Coelli and Arne Henningsen, *frontier: Stochastic frontier analysis*, 2008, R package version 0.9, `http://CRAN.R-project.org`.

[2] Yanqin Fan, Qi Li, and Alfons Weersink, *Semiparametric estimation of stochastic production frontier models*, Journal of Business and Economic Statistics **14** (1996), no. 4, 460–68.

[3] Tristen Hayfield and Jeffrey S. Racine, *Nonparametric econometrics: The np package*, Journal of Statistical Software **27** (2008), no. 5, 1–32.

[4] Paul W. Wilson, *FEAR: A software package for frontier efficiency analysis with R*, Socio-Economic Planning Sciences **42** (2008), no. 4, 247–254.

[1] Please note that the non-academic use of the **FEAR** package is restricted and that this closed-source software is available as binary package for MS-Windows only.