Quantile DEA: A Direct Linear Programming Based Approach to Obtaining Quantile Efficiency or Quantile Group Benchmarking Performance Estimates

Joseph Atwood
309E Linfield Hall
Montana State University
Bozeman, MT 59717 USA
Tel: 1-406-994-5614
Fax: 1-406-994-4838
jatwood@montana.edu

Saleem Shaik
Dept. of Agribusiness and Applied Economics
North Dakota State University
Fargo, ND USA

Selected Poster prepared for presentation at the 2016 Agricultural & Applied Economics Association Annual Meeting, Boston, Massachusetts, July 31-August 2

Copyright 2016 by J. Atwood, and S. Shaik. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.
Data envelopment analysis (DEA) was introduced by Charnes, Cooper and Rhodes (1978) in operational research and popularized in a more informative and easily applied way by Fare et al. (1994). Since its introduction, the number of applications of DEA to firms, industry, state and countries has exploded. Along with the applications, researchers are beginning to identify and address issues associated with the empirical application of DEA including observational outliers and robust estimators.

In this research, a new robust procedure with good statistical properties is developed using a linear lower partial moment stochastic inequality. The procedure works deviations from constraints in such a way that guarantees that no more than q x 100 percent of points lie external to the hull.

The new procedure is “Quantile DEA (qDEA)” and simultaneously identify the constraints to relax in a second stage conventional DEA model. Similar but less efficient procedures date as far back as 1977. Charnes et al. (1977) introduced Partial Moment Inequality: Atwood (1985) generalized Berk-Hihn (1982) and Atwood (1985). Atwood’s linear PM inequality can be utilized to implement chance constraints within a conventional LP model.

LP duality theory shows that the qDEA Stage 1 efficiency scores are projected from both the desired support points and the external points. Stage 2 in qDEA relaxes the DEA LP constraints for the stage 1 external points. The qDEA stage 2 program generates the desired projections and quantile efficiency distances and efficiencies. The following table presents estimated efficiency scores for a slightly modified problem from Cooper, Seldford, and Tone (2007). The following table presents the estimated efficiency scores with 0, 1, and 2 external points.

Under a reasonable set of assumptions, subsample or nCM bootstrapping can be used to estimate the significance, bias corrections, and confidence intervals of parameter estimates. (Politis et. al. (1999, 2001) ; Simar and Wilson (2011); Geyer (2015). With the nCM bootstrap B subsamples (X, L) of size m < n are sampled from the original input-output sample (X, L) and used to construct a simulated set of parameter realizations: 

\[ \hat{\theta}_m = \frac{1}{B} \sum_{b=1}^{B} \hat{\theta}_{m,b} \] 

The mean or median of the B resulting values can be used as a bias corrected parameter estimate. The B values \( \hat{\theta}_m \) are also bias corrected and their quantile values can be used to construct confidence intervals. A complication arises in that there is no known way to select the appropriate level for m and common practice is to estimate a set of \( \hat{\theta}_m \) values for multiple m values and use suggestions by Politis et al. (2001) and Simar and Wilson to select an appropriate level for m. Computation time was reduced substantially by using only five m values in the set mlist where 

\[ \text{mlist} = \text{floor}(\exp(seq(log(sqrt(n)),log(sqrt(n)),length.out = 5)) \]

For example, when n = 1000, the set of m values is (31,47,70,105, 158).

A large number of Monte Carlo simulations were run for various sample sizes and data generating processes. Empirical coverage levels were then computed as the proportion of times that the estimated confidence interval contained the population parameter. 95% qDEA coverage levels similar to those obtained by Simar and Wilson for the DEA model were obtained.