3rd International Workshop on Empirical Methods in Energy Economics (EMEE2010)

Surrey Energy Economics Centre (SEEC)

University of Surrey, UK

24th – 25th June 2010

NOTE:

The following Presentation represents *Work in Progress* for discussion at the EMEE2010 workshop. It therefore must not be referred to without the consent of the author(s).

Sponsored by:

Overcoming Data Limitations in Nonparametric Benchmarking: Applying PCA-DEA to Natural Gas Transmission

Maria Nieswand, Astrid Cullmann, Anne Neumann

Poster Presentation at EMEE 2010,
June, 24-25 2010, University of Surrey, Guildford, GB

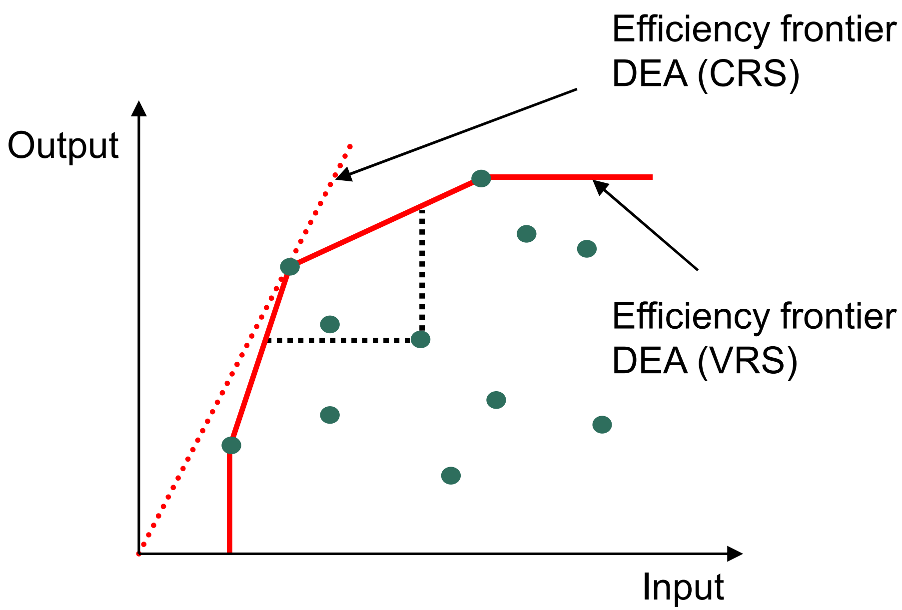
Motivation

- Benchmarking is a powerful tool in regulation of deregulated network industries, e.g., natural gas transmission
- Data limitations due to:
  - Limited number of companies (because of former monopolistic market structures)
  - Yearly conducted analysis
- Regulatory practice often rely on nonparametric efficiency measurement (Data Envelopment Analysis)
  - Require large datasets in order to derive meaningful results
  - Suffer from bad ratios of included variables over the number of observations (*curse of dimensionality*)

Issue: Limitation in data affects nonparametric efficiency measurement

- Noticeable part of regulation in Europe relies on nonparametric analysis, i.e. Data Envelopment Analysis (DEA)

Figure: Concept Data Envelopment Analysis



- Desirable: many variables → many dimensions
- But only few observations
- Lack of discriminatory power of DEA ("curse of dimensionality)
- Over-estimation (inefficient companies are incorrectly defined as efficient)
- Uncovered cost reduction potential

Purpose of the paper

- Provide a pragmatic approach for European regulators who predominantly undertake efforts for national benchmarking

- Derive a benchmarking model that:
  - Can deal with a high number of variables
  - Can deal with few observations at the same time
  - Does not reduce discriminatory power substantially

- Combining DEA with principle components analysis (PCA)
  - “PCA-DEA” (Adler and Yazhemsky, 2009)

► Reducing dimensions and keeping most of information at the same time
► Increasing rate of convergence
► Overcome problem of data limitation

Piecewise Linear Programming

Traditional DEA approach

$$\min_{\theta,\lambda} \theta$$

$$s.t.\ Y\lambda - s_Y = Y_j$$

$$-X\lambda - s_X = \theta X_j$$

$$e\lambda = 1$$

$$\theta, \lambda, s_Y, s_X \geq 0$$

PCA- DEA approach

$$\min_{\theta,\lambda} \theta$$

$$s.t.\ Y_{PC}\lambda - L_Y s_{PC} = Y_{PC,j}$$

$$-X\lambda - s_X = \theta X_j$$

$$L_Y^{-1} Y_{PC} \geq s_{PC}$$

$$e\lambda = 1$$

$$\theta, \lambda, s_{PC}, s_X \geq 0$$

Model Specification

Figure: "OPEX-Benchmarking"

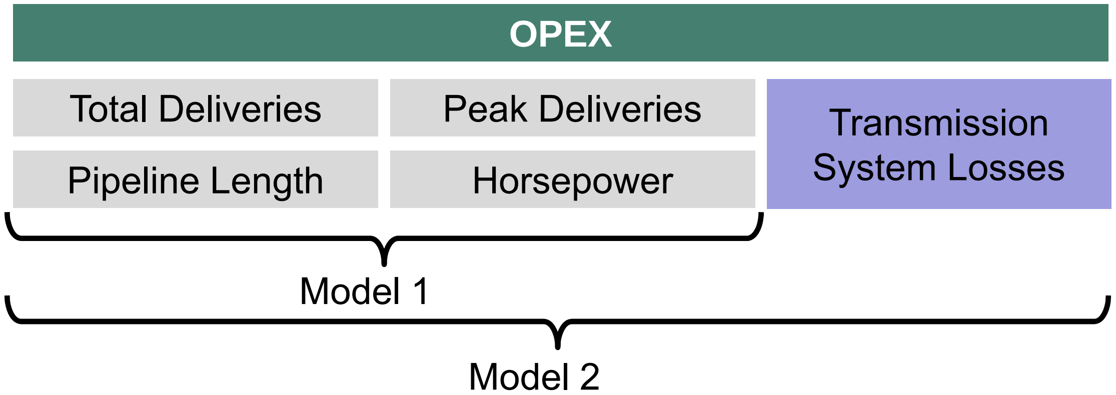


Table: Model specification

	Model 1		Model 2	
	DEA	PCA-DEA	DEA	PCA-DEA
DEA	x		x	
PCA-DEA		x		x
VRS	x	x	x	x

Note: x denotes the presence of the assumption in each model specification.

Results PCA

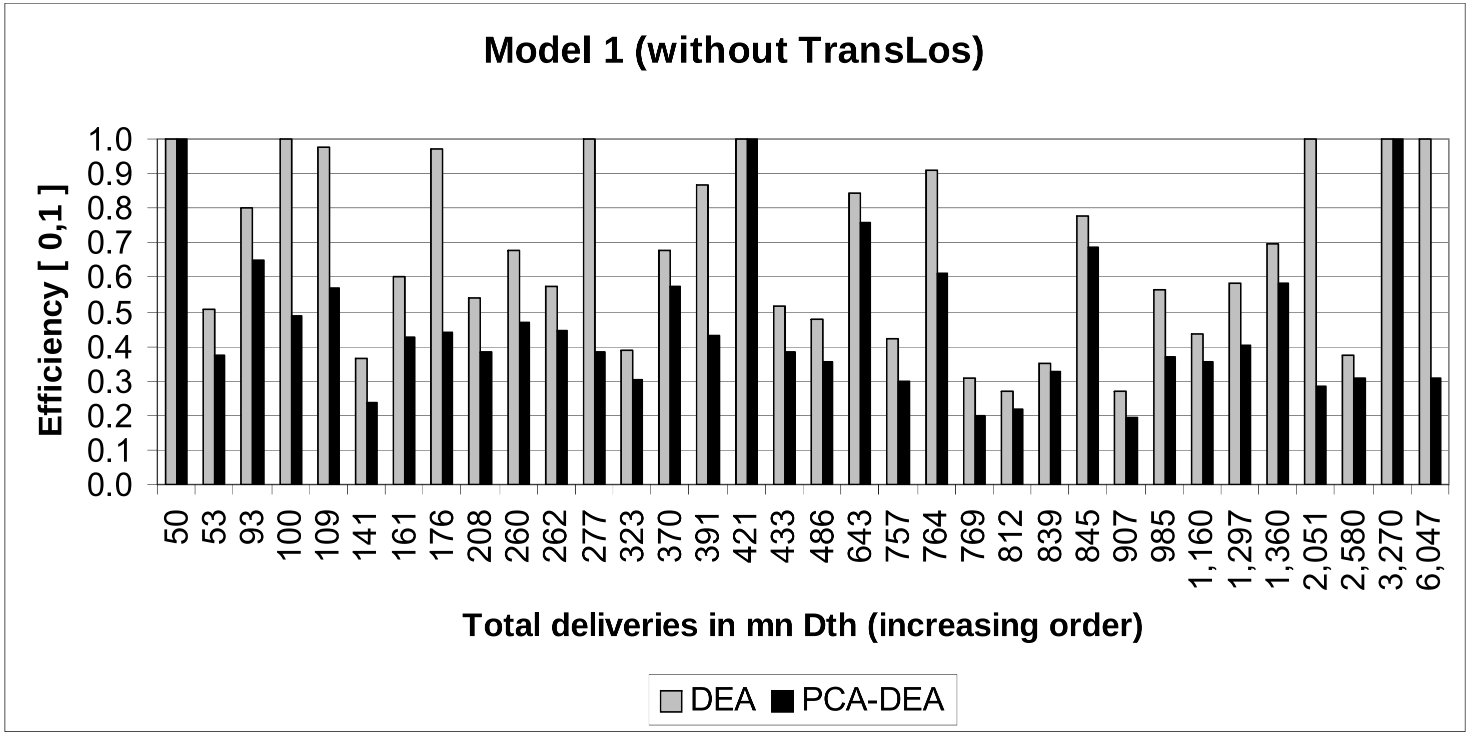
Table: Principle Components Analysis

	Model 1 (without TransLos) Variance explained (%)			Model 2 (with TransLos) Variance explained (%)		
	Inputs	Outputs	Cumulative	Inputs	Outputs	Cumulative
PC 1	100	87.76	87.76	100	82.19	82.19
PC 2		7.52	95.28		8.34	90.53
PC 3		3.35	98.62		5.71	96.24
PC 4		1.38	100		2.68	98.92
PC 5					1.08	100

- We chose the first two principle components (PC 1 and PC 2)
- Dimensions decrease from four (in Model 1) and five (in Model 2) to two

DIW BERLIN

Results of Efficiency Measurement I

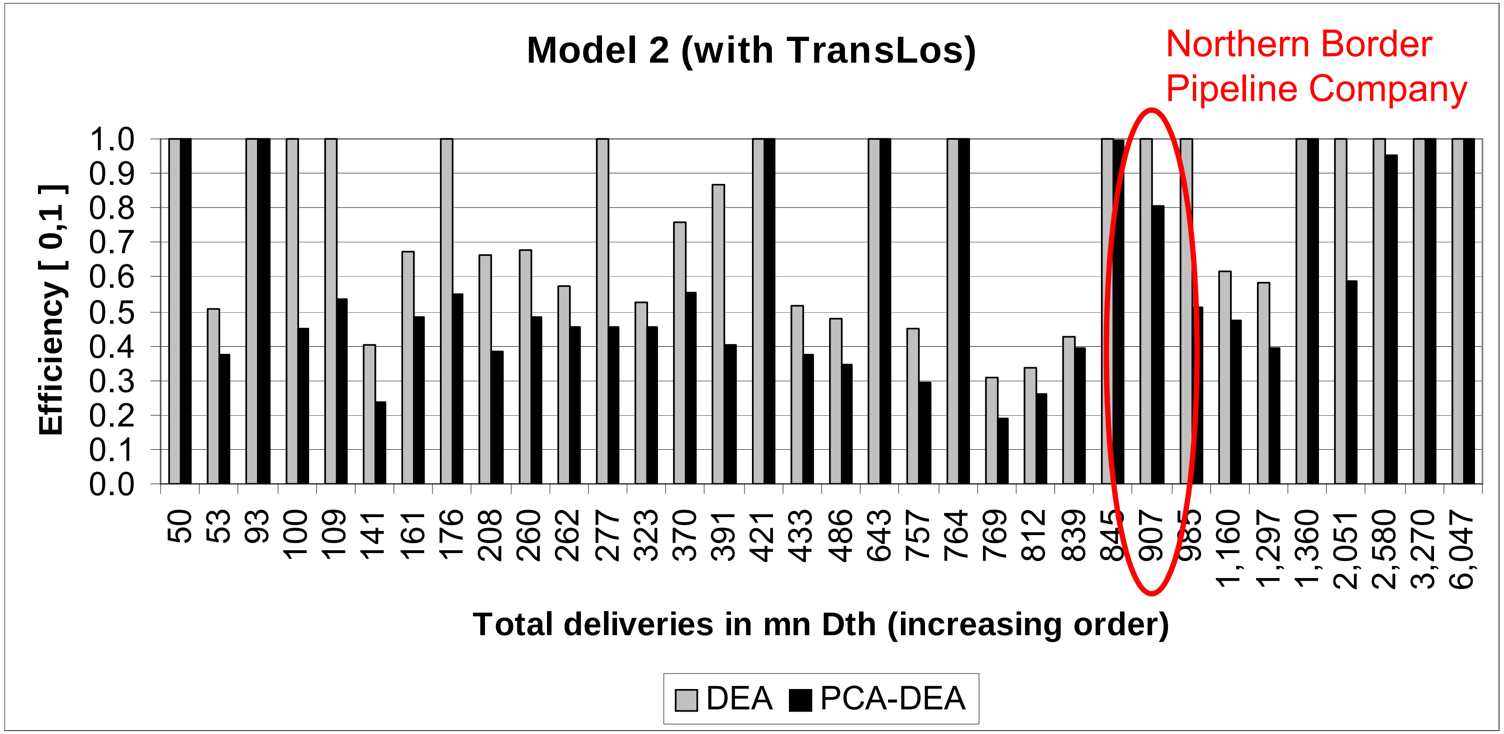


► Lower efficiency lower under PCA-DEA (mean 66.89%) than under DEA (46.54%) methodology

► Probability of over-estimation decreases under PCA-DEA

► The more variables the more companies defined as efficient

Results of Efficiency Measurement II



► Higher efficiency under Model 2 compared to Model 1 (Note: dimensions under PCA-DEA specifications are the same!)

► Impact of specialists lower

DIW BERLIN

Results of Efficiency Measurement III

Figure: Peers of Northern Border Pipeline Company in Model 1 and Model 2 under PCA-DEA specification

Pipeline / Variable	OPEX	Total deliveries	Pipeline Length	Peak deliveries	Installed horsepower	Transmission system losses
Unit	mn USD	mn Dth	Miles	mn Dth	thou HP	thou Dth
Northern Border	165.3	907.0	1,399	2.6	536.6	77.9
Peers in Model 1						
IPOC	9.3	420.6	414	1.4	78.3	489.4
Transcont	117.3	3270.0	10,325	8.4	1434.3	6684.6
Peers in Model 2						
Dominion	70.7	1360.1	3,344	4.0	350.2	398.5
El Paso	373.4	6046.7	10,240	5.1	1136.4	3038.8

Note: Northern Border = Northern Border Pipeline Company; IPOC = IPOC as Agent/Iroquois Gas Trans. Sys. L.P.; Transcont = Transcontinental; Gas Pipe Line Corporation; Dominion = Dominion Transmission, Inc.; El Paso = El Paso Natural Gas Company

- ► PCA-DEA can change the reference set (peers)
- ► PCA-DEA and the inclusion of relevant variables yield in a reference set that is structurally similar to the company under consideration
- ► Identifying “local” conditions key feature of DEA

Conclusion

- (Nonparametric) Benchmarking important tool for regulators
- Practical data limitations decrease discriminatory power of nonparametric efficiency measurement
- Apply method that deals with many variables and few observations at the same time
- PCA-DEA improves nonparametric efficiency measurement with small sample sizes and relatively many variables
- Extensive importance for regulator and companies (mitigation of conflict between number of observations and variables, and “local” conditions)

Thank you for your attention.

Questions and comments are welcome!

Contact:

Maria Nieswand, DIW Berlin

mnieswand@diw.de

Selected References

- Adler, N., & Golany, B. (2001). Evaluation of deregulated airline networks using data envelopment analysis combined with principle component analysis with an application to western europe. European Journal of Operational Research, 132 , 260-73.
- Adler, N., & Golany, B. (2002). Including principle component weights to improve discrimination in data envelopment analysis. Journal of Operational Research Society, 53 , 985-91.
- Adler, N., & Yazhemsky, E. (2009). Improving discrimination in data envelopment analysis: Pca-dea versus variable reduction. which method at what cost? European Journal of Operational Research, forthcoming, doi:10.1016/j.ejor.2009.03.050.
- CEER (2006). International benchmarking and regulation of European gas transmission utilities. Available from http://www.energy-regulators.eu
- EC (2009). Report on progress in creating the internal gas and electricity market. Technical annex.
- Haerdle, W., & Simar, L. (2003). Applied multivariate statistical analysis.
- Jamasb, T., Pollitt, M. G., & Triebs, T. (2008). Productivity and efficiency of us gas transmission companies: A European regulatory perspective. Energy Policy, 36, 3398-412.
- Pastor, J. T. (1996). Translation invariance in data envelopment analysis: A generalization. Annals of Operations Research, 66 , 93-102.
- Simar, L., & Wilson, P. W. (2008). Statistical inference in nonparametric frontier models: Recent developments and perspectives.
- Ueda, T., & Hoshiai, Y. (1997). Application of principle component analysis for parasimonious summarization of dea inputs and/ or outputs. Journal of the Operations Reseach, Society of Japan, 40 (4), 466-78.