Deutsches Institut für Wirtschaftsforschung

DIW BERLIN

Overcoming Data Limitations in Nonparametric Efficiency Measurement

A PCA-DEA Application to Natural Gas Transmission

Maria Nieswand, Astrid Cullmann, Anne Neumann

INFRADAY 2009
October, 10 2009, Berlin, Germany

Agenda

Motivation

Methodology

Data

Results

Conclusion

Motivation

- 1st step **WAS** restructuring international electricity and natural gas markets towards deregulation

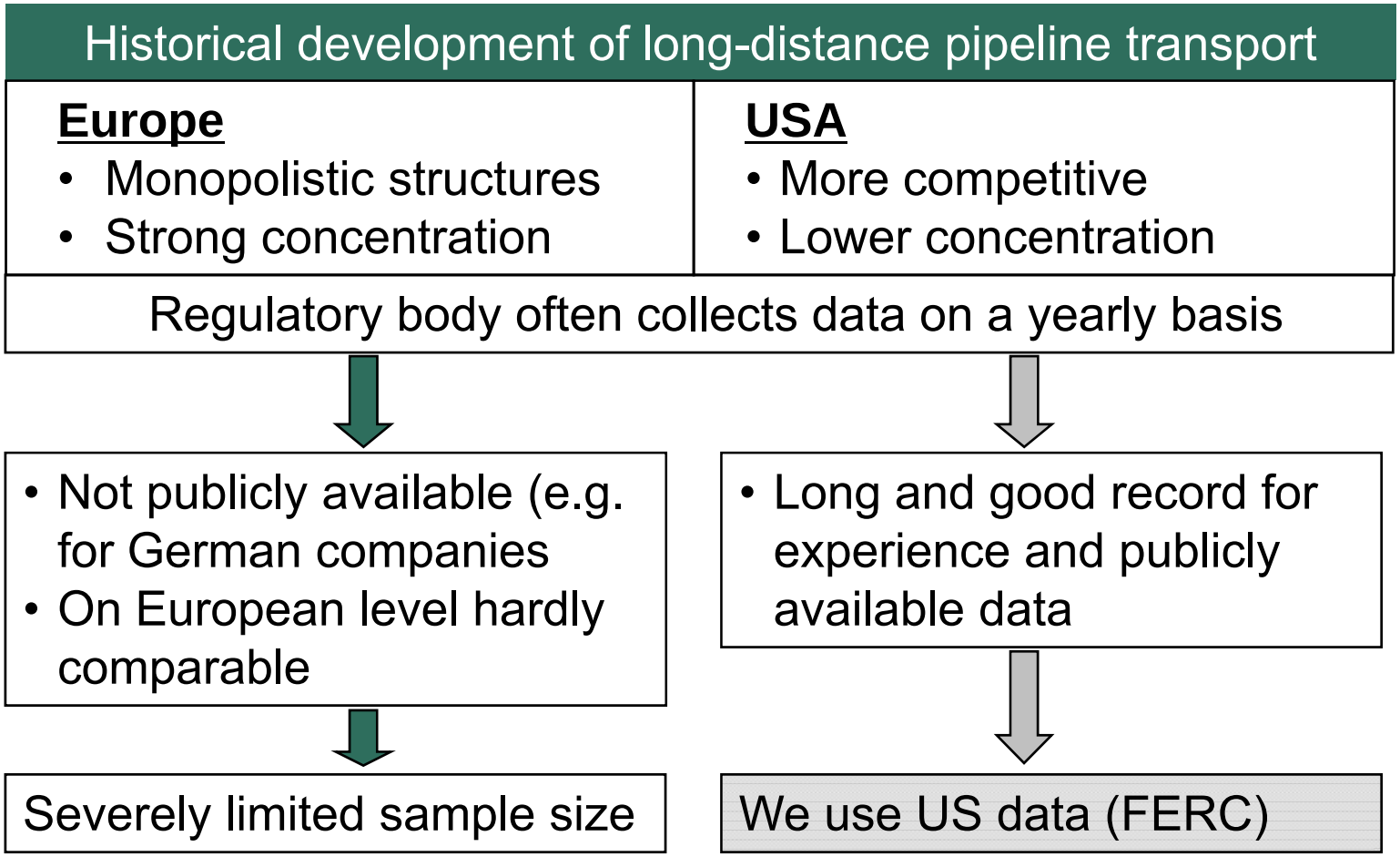
- 2nd step **IS** benchmarking in deregulated network industries

  ► Benchmarking remains a powerful tool
  ► Determine the most efficient way of production and supply

- Regulation is worldwide based on both, parametric and nonparametric benchmarking techniques (CEER, 2006; Haney and Pollitt, 2009)

  ► Differ in requirements on underlying sample size in order to derive meaningful results (relative efficiency)

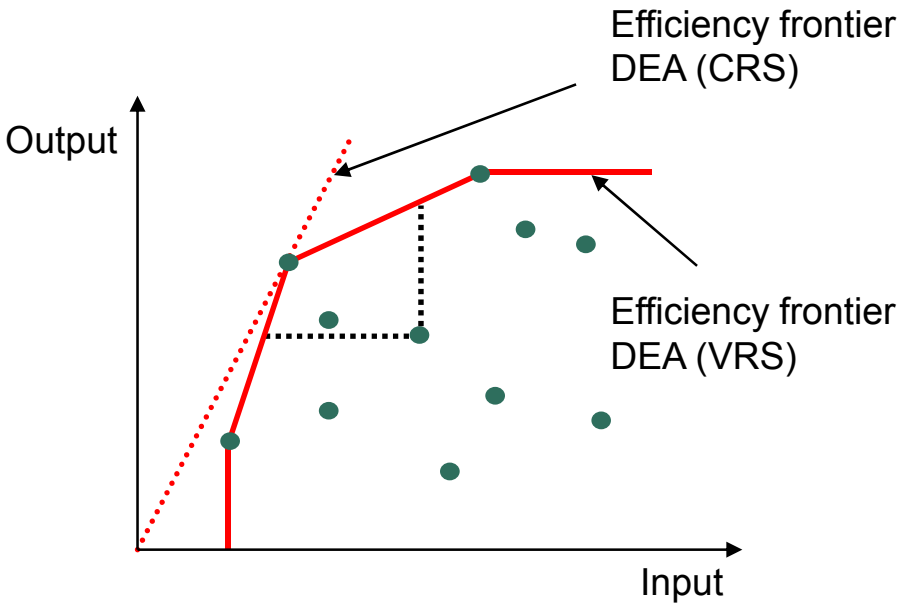
Practically Limited Data Amount in Europe



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

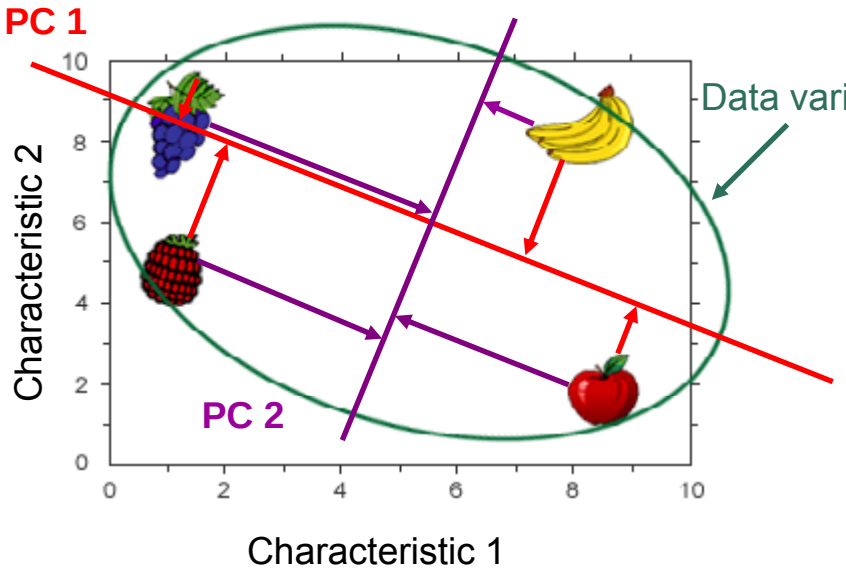
- Provide empirical demonstration for practical approach of efficiency measurement

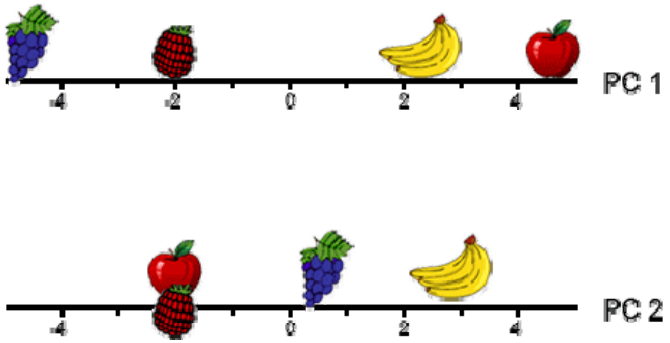
- Derive a benchmarking model that:
  - Can deal with a high number of variables
  - Can deal with few observations at the same time
  - Does not reduce discriminatory power substantially

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

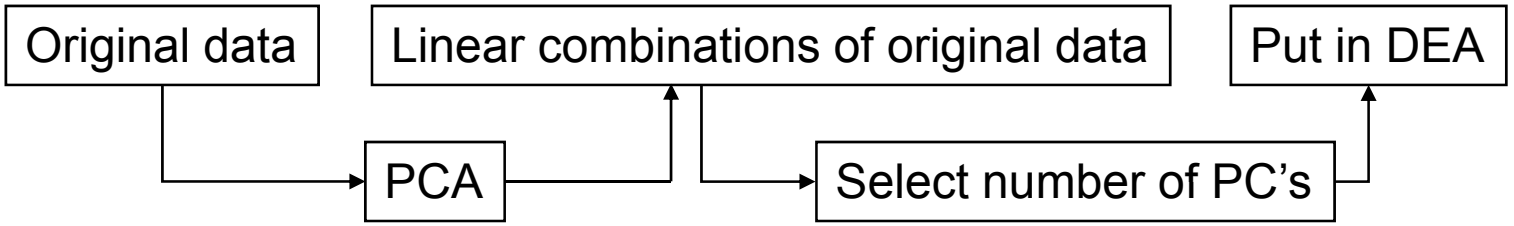
    ► Reducing dimensions and keeping most of information at the same time
    ► Increasing rate of convergence
    ► Overcome problem of data limitation

Principle of PCA-DEA

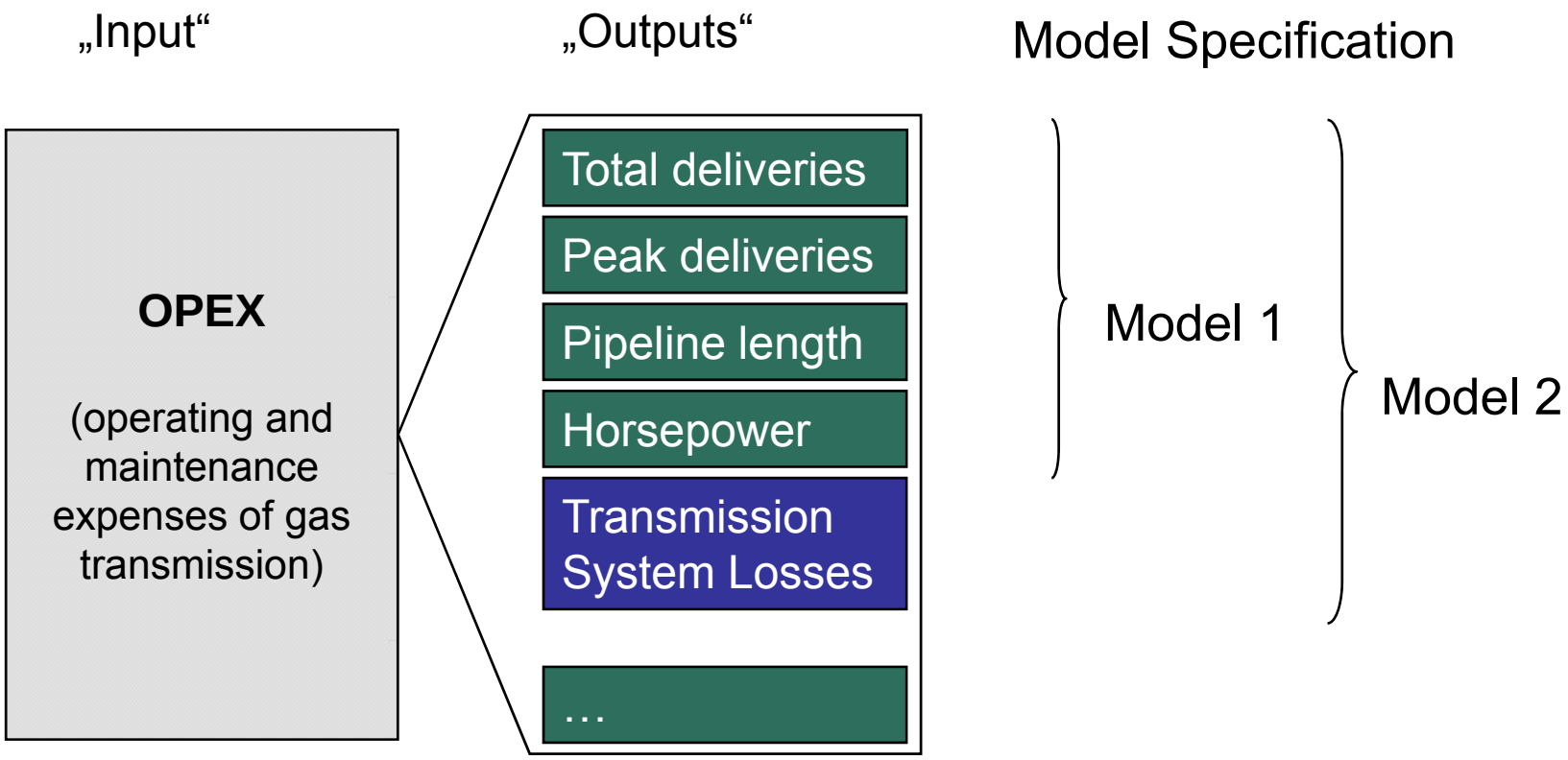




Source: http://www.chemgapedia.de/



Costs and Cost Drivers of NGT Companies



US Natural Gas Transmission Pipelines 2007, 37 observations

Table: Descriptive Statistics of US Natural Gas Transmission Pipelines 2007 (onshore)

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
Sum	2,860.32	34,191.24	127,783.20	86.81	11,003.22	38,677.68
Minimum	1.25	49.93	59.00	0.19	9.00	0.00
Maximum	402.67	6,046.71	14,463.20	8.44	1,434.27	6,684.60
Median	31.50	432.91	1,680.40	1.68	125.95	615.66
Mean	77.31	924.09	3,453.60	2.35	297.38	1,045.34
Std. dev.	99.61	1,255.53	3,703.33	2.12	371.72	1,399.32

Source: FERC Form No. 2

- DEA is sensitive to outliers
  - ► Outlier detection based on super-efficiency
  - ► Sample shrinks to 34 observations

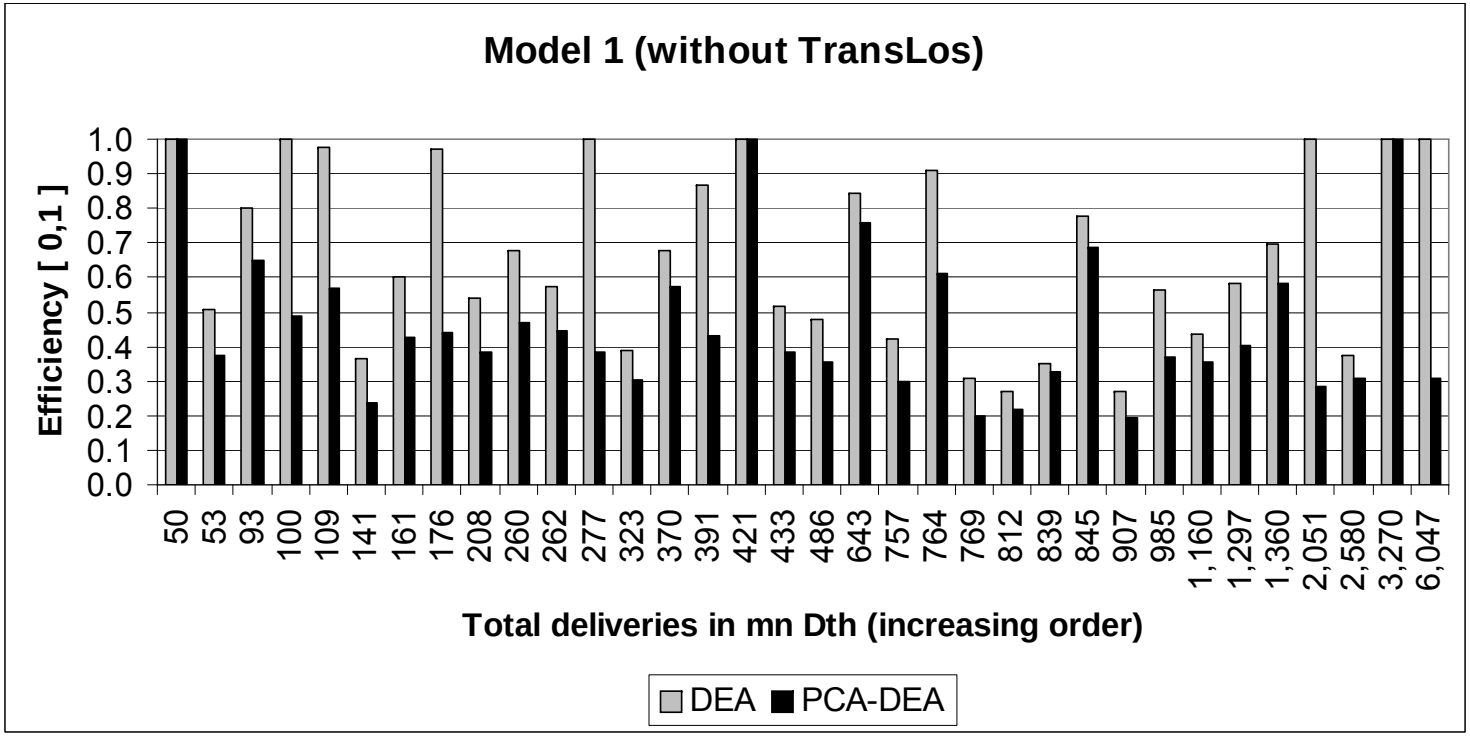
Results PCA

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

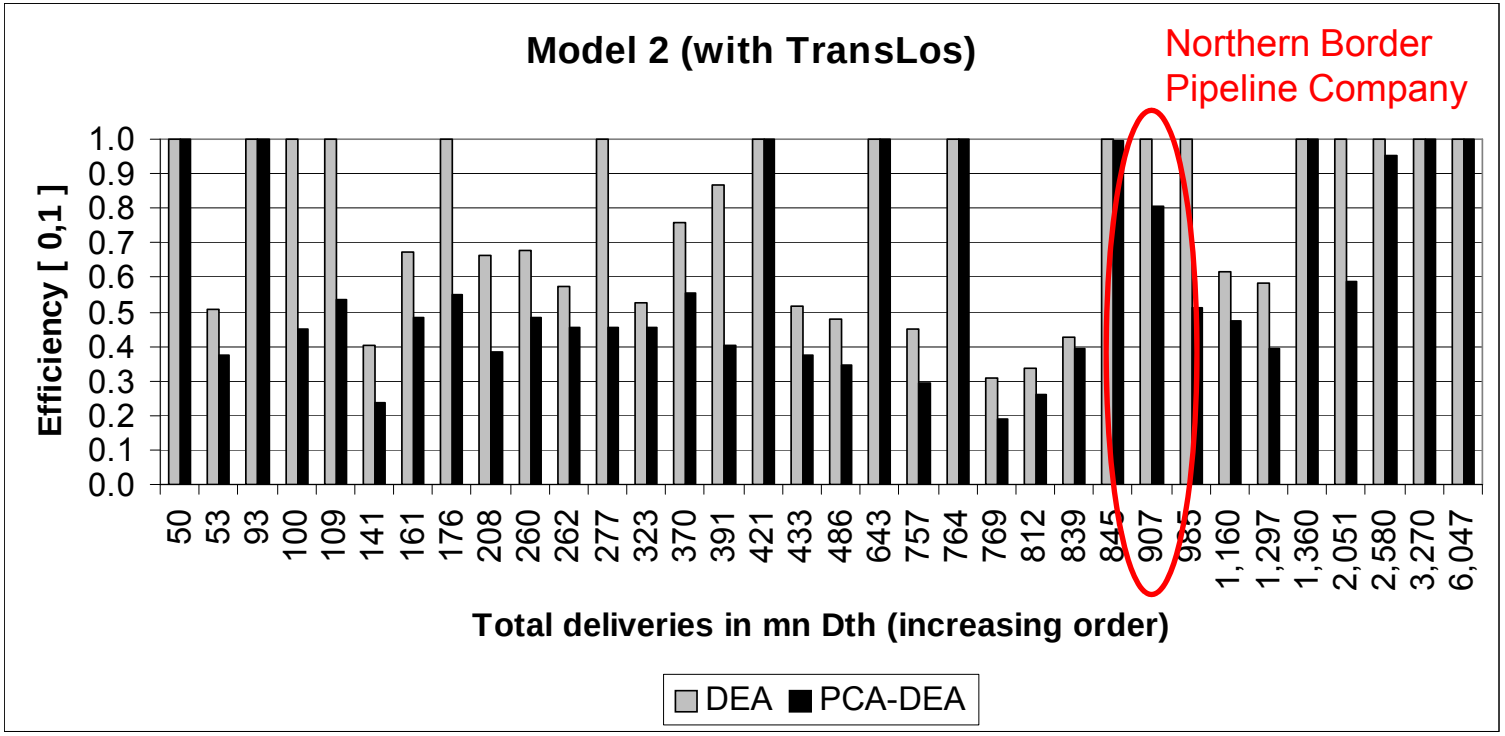
- We chose the first two principle components (PC 1 and PC 2)
- Dimensions decrease from four (in Model 1) and five (in Model 2) to two

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

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.

Backup

Natural Gas Transmission Grid in Europe

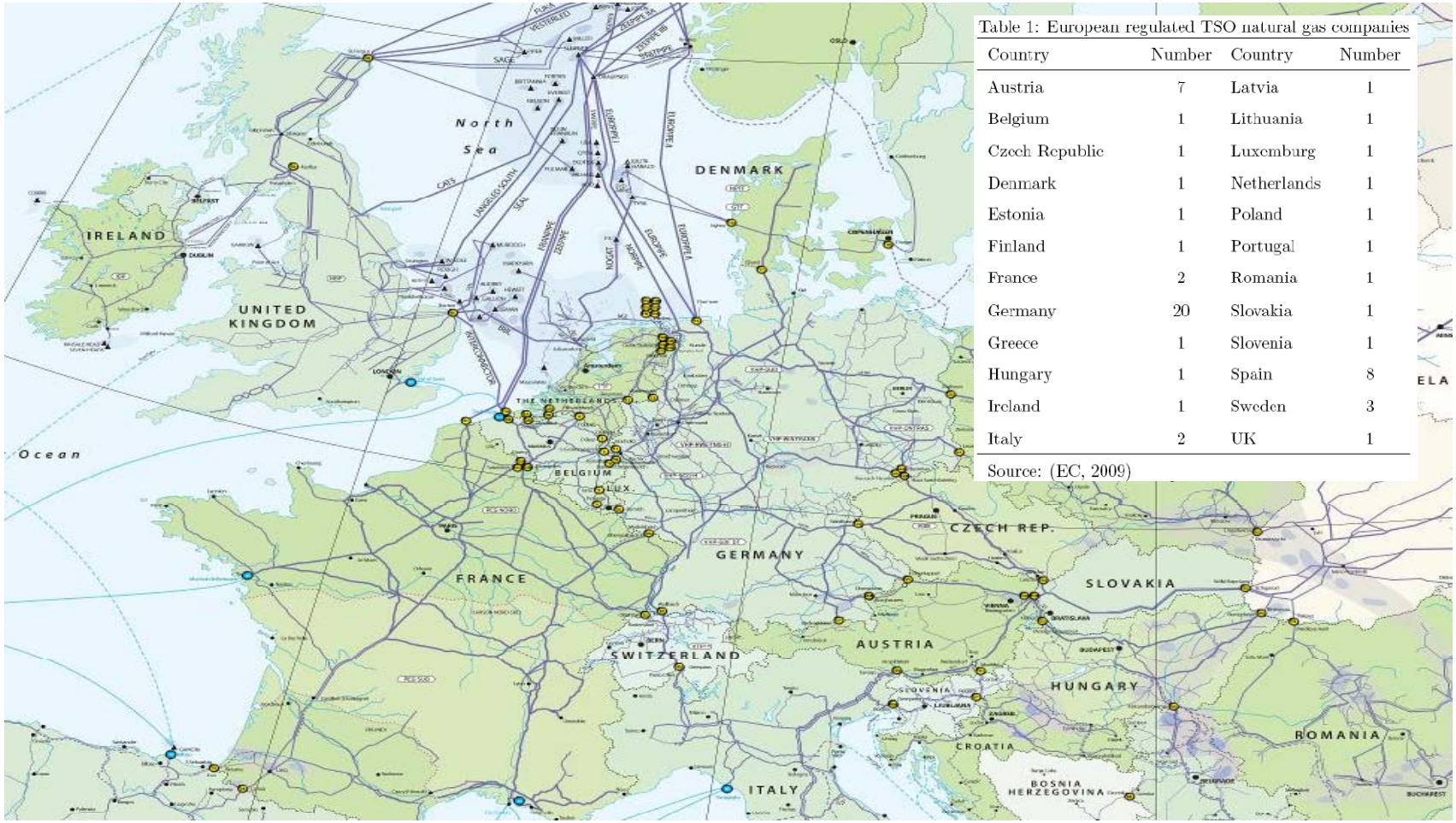


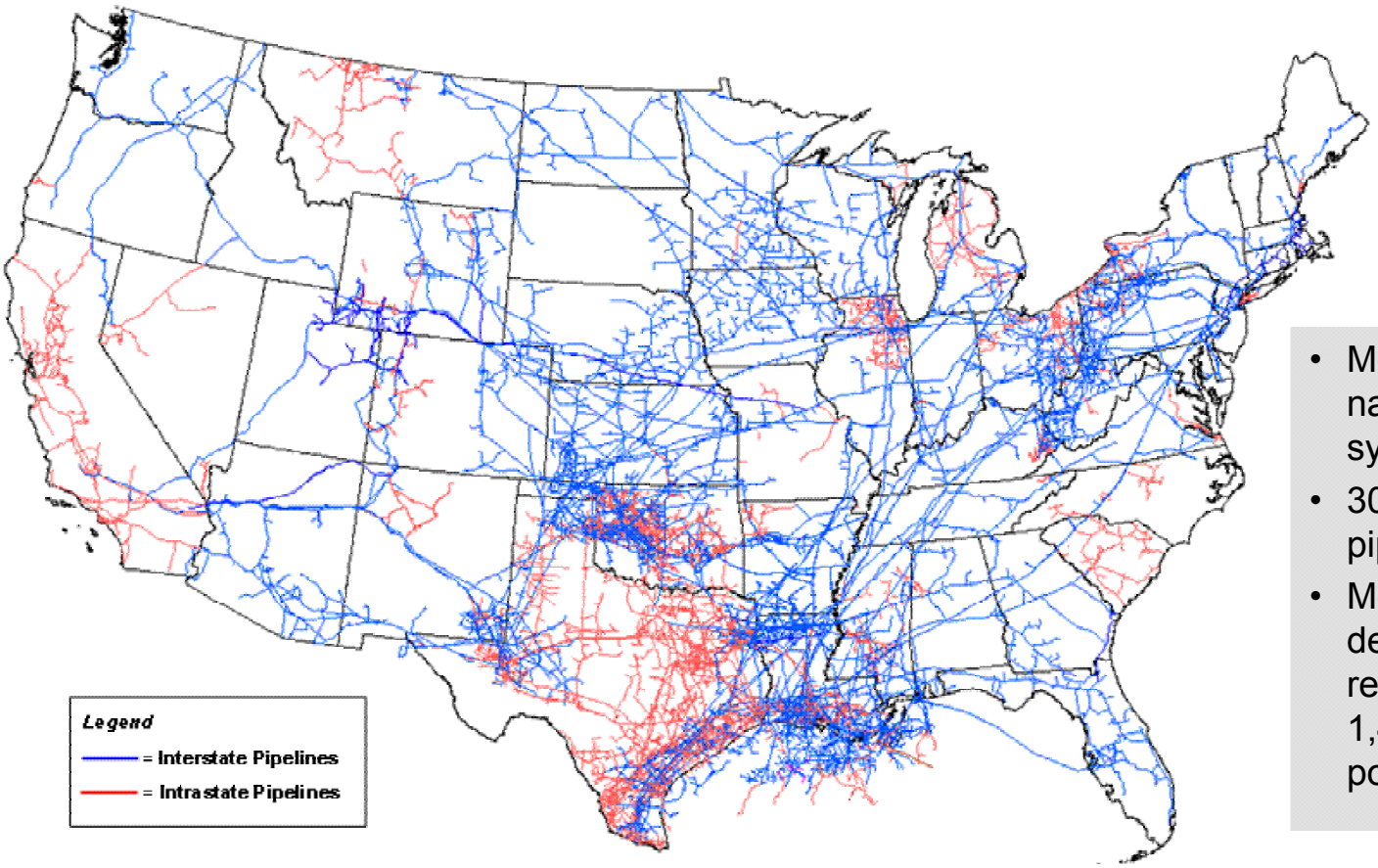
Table 1: European regulated TSO natural gas companies

Country	Number	Country	Number
Austria	7	Latvia	1
Belgium	1	Lithuania	1
Czech Republic	1	Luxemburg	1
Denmark	1	Netherlands	1
Estonia	1	Poland	1
Finland	1	Portugal	1
France	2	Romania	1
Germany	20	Slovakia	1
Greece	1	Slovenia	1
Hungary	1	Spain	8
Ireland	1	Sweden	3
Italy	2	UK	1

Source: (EC, 2009)

Source: GTE (Gas Transmission Europe)

Natural Gas Transmission Grid in USA



Source: Energy Information Administration, Office of Oil & Gas, Natural Gas Division, Gas Transportation Information System

- More than 120 natural gas pipeline systems
- 302,000 miles of pipelines
- More than 11,000 delivery points; 5,000 receipt points, and 1,400 interconnection points

Piecewise linear Programming

Traditional DEA approach

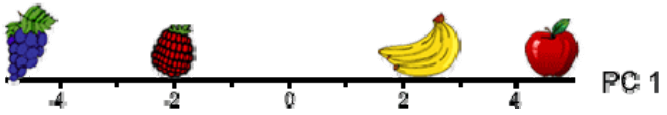
$$\min_{\theta,\lambda} \ \theta$$
$$\text{s.t. } Y\lambda - s_Y = Y_j$$
$$-X\lambda - s_X = \theta X_j$$
$$e\lambda = 1$$
$$\lambda, \theta, s_Y, s_X \geq 0$$

PCA- DEA approach

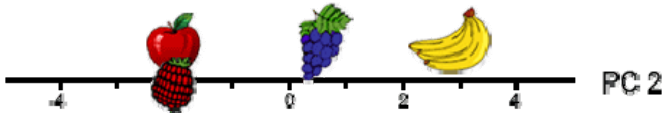
$$\min_{\theta,\lambda} \ \theta$$
$$\text{s.t. } Y_{PC}\lambda - L_Y s_{PC} = Y_{PC_j}$$
$$-X\lambda - s_X = \theta X_j$$
$$L_Y^{-1} Y_{PC} \geq s_{PC}$$
$$e\lambda = 1$$
$$\lambda, \theta, s_{PC}, s_X \geq 0$$

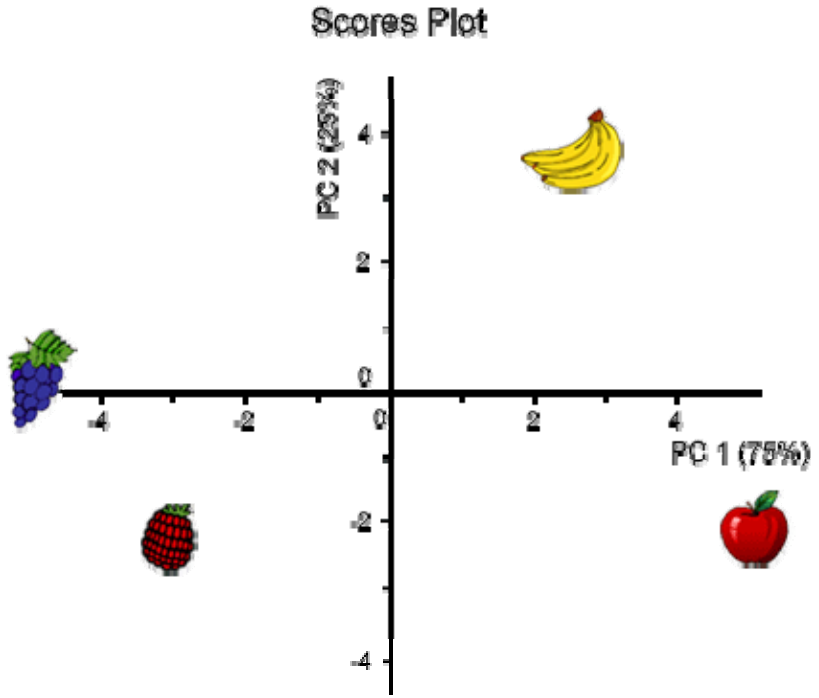
Principle Components Analysis

Distinguish fruit varieties



Distinguish colors





Source: http://www.chemgapedia.de/