

Regulatory benchmarking: applicability and effectiveness

Per J. AGRELL
UC Louvain / CORE

Bundesnetzagentur “Efficiency Benchmarking”
2024-09-24

Outline

Regulation and efficiency analysis

- › Acceptability and adequacy

Regulatory benchmarking

- › Criteria

Heterogeneity

Process

Outliers

Conclusions

Example

- › International gas models

Université catholique
de Louvain

UCL

REGULATION AND EFFICIENCY ANALYSIS

Some principles

The relevant model of competition to inform regulatory policy is not one of atomistic or perfect competition, but rather one that evaluates and rewards the performance of regulated entities.

Weisman and Pfeifenberger, 2003

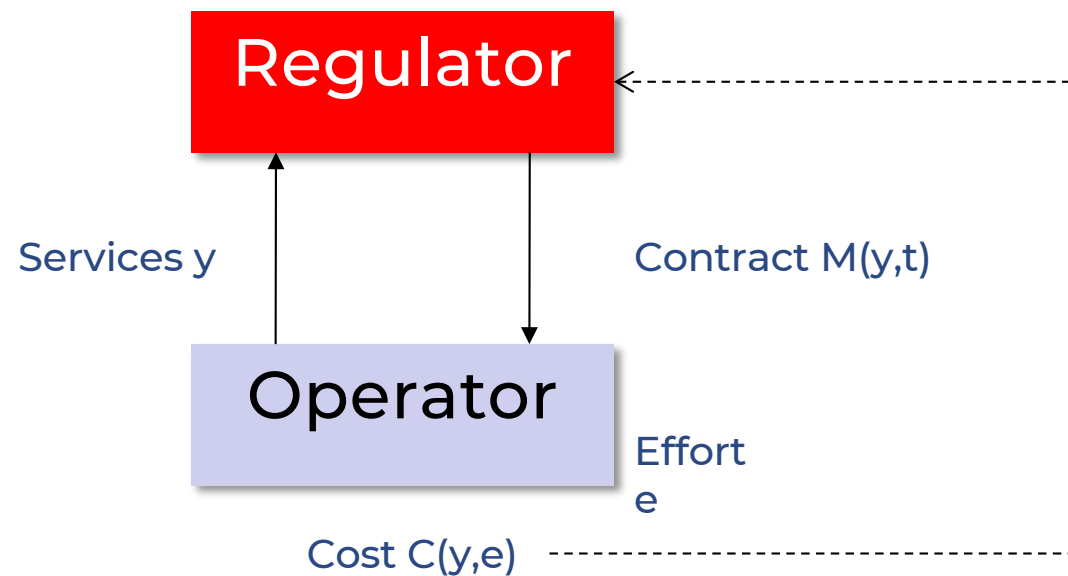
Acceptability and adequacy of regulatory benchmarking

***Acceptable** is a method compliant with current national and European laws and directives.*

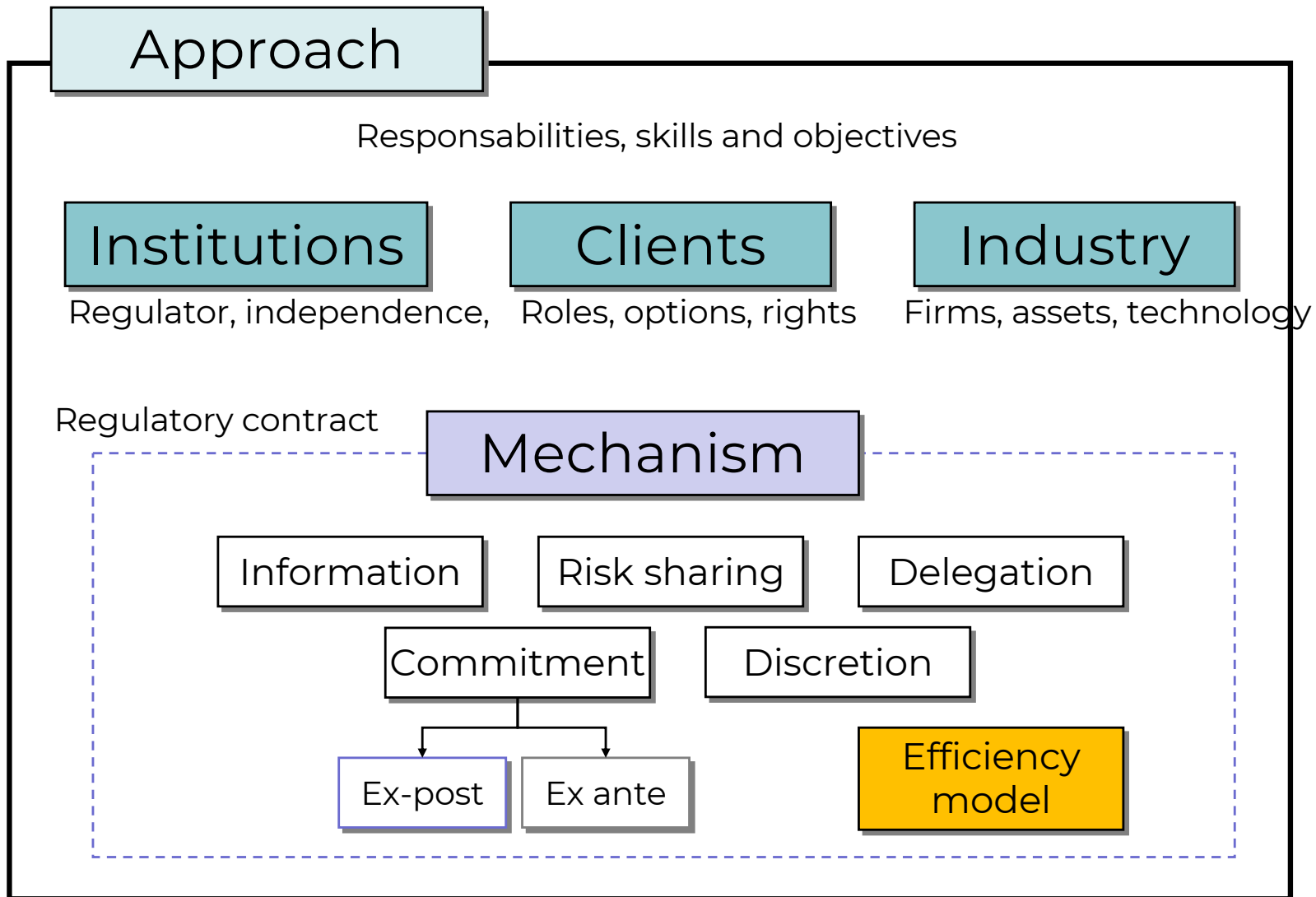
***Adequate** is a method fit for achieving higher economic and technical objectives in its application in energy network regulation.*

INCENTIVE REGULATION

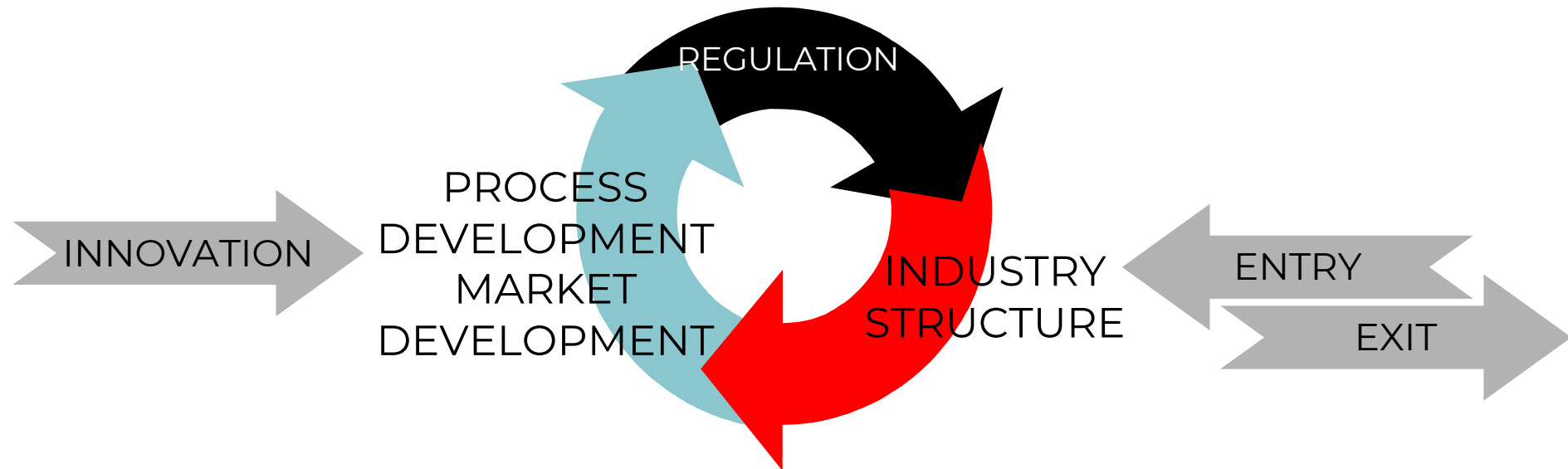
Conventional view



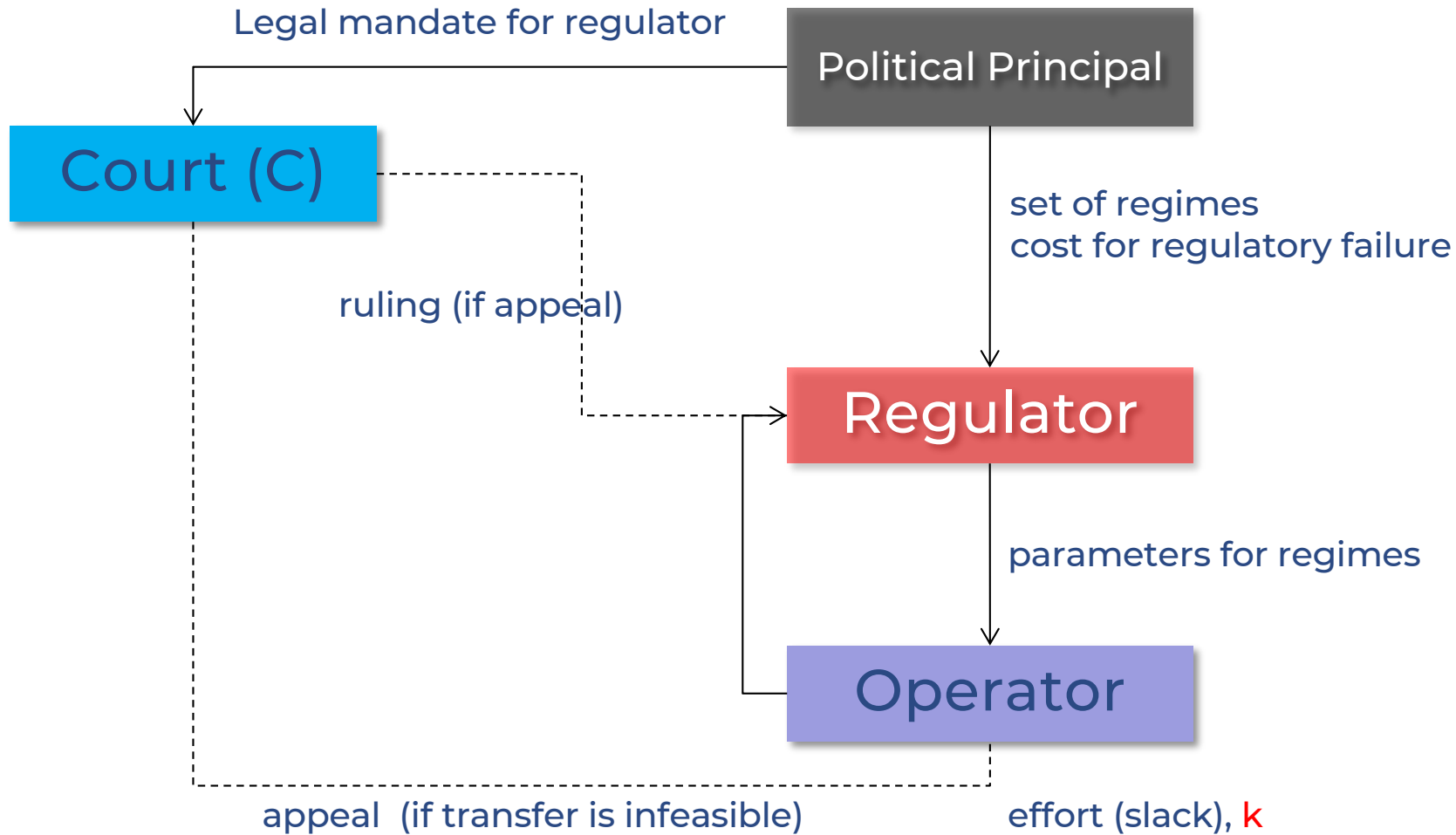
Infrastructure access, unbundled firm, inelastic demand for service
Cost is observable and verifiable, effort is unobservable, multi-output service provision
High-powered regulation is optimal: Laffont (1994), et al.
Practical implementations: yardstick regimes: Schleifer (1985), Laffont and Tirole (1986)



Endogenous structural effects



Is there an implicit sector agenda?



Common benchmarking methods

	Deterministic	Stochastic
Parametric	COLS	SFA
Non-parametric	DEA	RDEA (Simar-Wilson) StonED
Normative	Reference models	Simulation models

Concerns for model use in regulation

Method validation

- › Sound theoretical basis for policy making

Data validation

- › Procedural impact

Controllability

- › Inputs under management control
- › Environmental factors strictly exogenous

Equity

- › Process independent and unbiased in size, governance, etc

REGULATORY BENCHMARKING

Efficiency in regulated business and regulation using efficiency

Prospective (academic) studies

- › Ex post choice of parameters
- › Ad hoc definition of reference set
- › Category-based effects (private/public, reform, governance)
- › Choice of model depends on exact question

Regulatory benchmarking

- › Ex ante choice of model, based on past data
- › Formal definition and application of outliers
- › Direct use in the setting of standards, tariffs or revenues

Requirements for regulatory cost norms

A cost norm for regulatory use must respect

- › **Feasibility** (evidence based estimate)
- › **Neutrality** (unbiased estimate)
- › **Robustness** (to data errors)
- › **Repeatability** (endogeneity of parameters)

Model objective

- › The cost norm is based on **structural comparability**
- › The resulting rulings should be **robust to judicial recourse**

R1. Feasibility

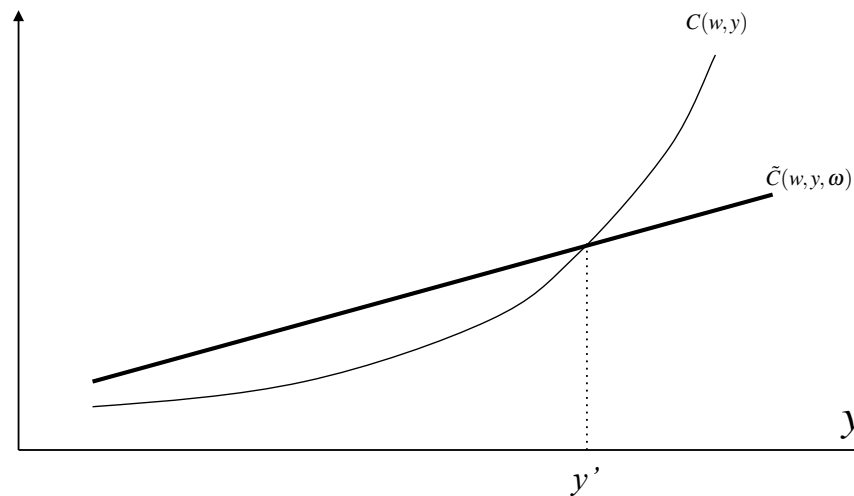


Figure 1: Feasibility violation of cost norm $\tilde{C}(w,y|\omega)$ from scale y' .

1. *Feasibility*: all estimates in $\tilde{C}(w,y,\omega)$ are realizable for an operator.

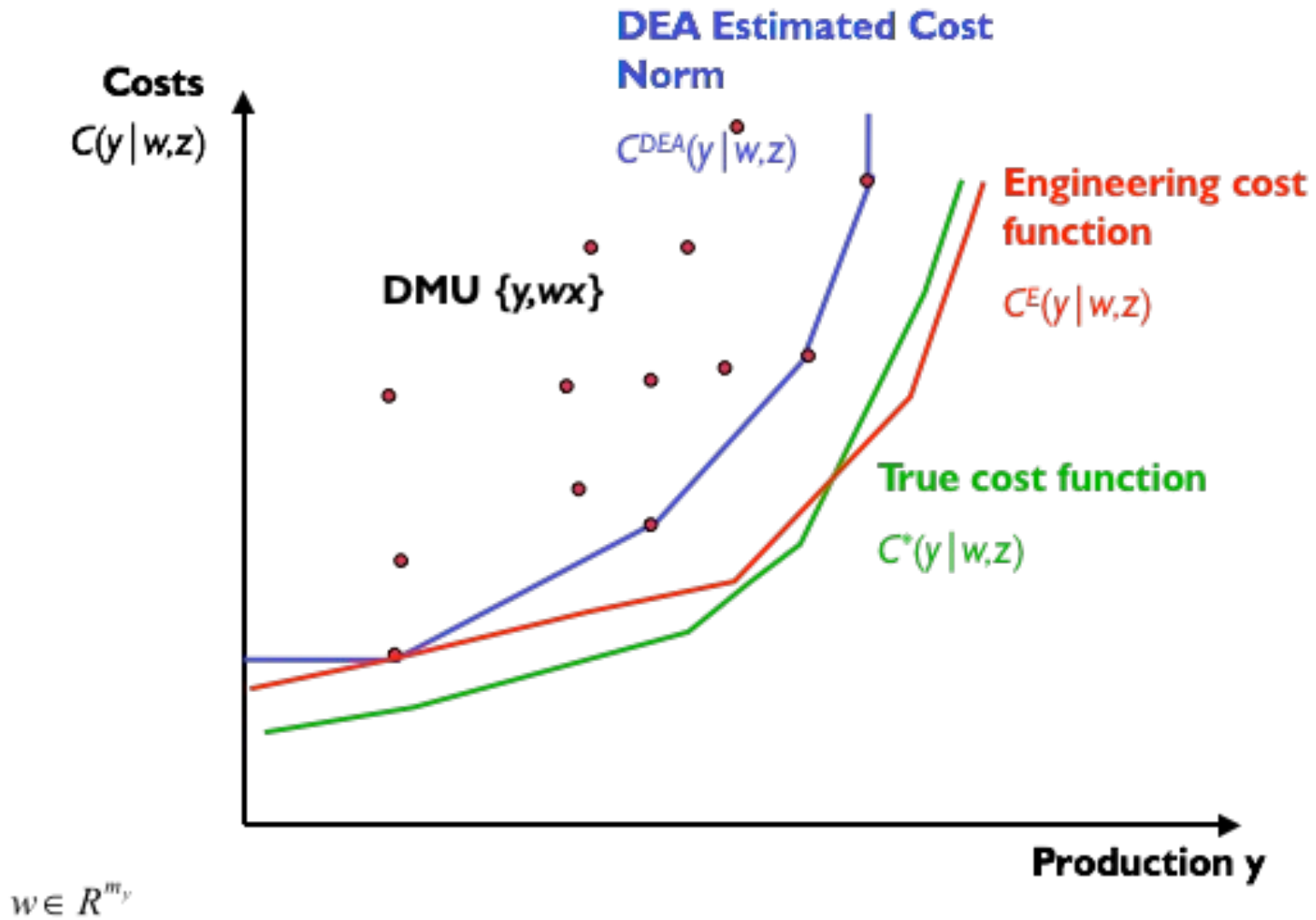
$$E[\tilde{C}(w,y|\omega) - C(w,y)] \geq 0 \text{ for all } w,y \text{ in } \omega.$$

R1 Feasibility evaluations

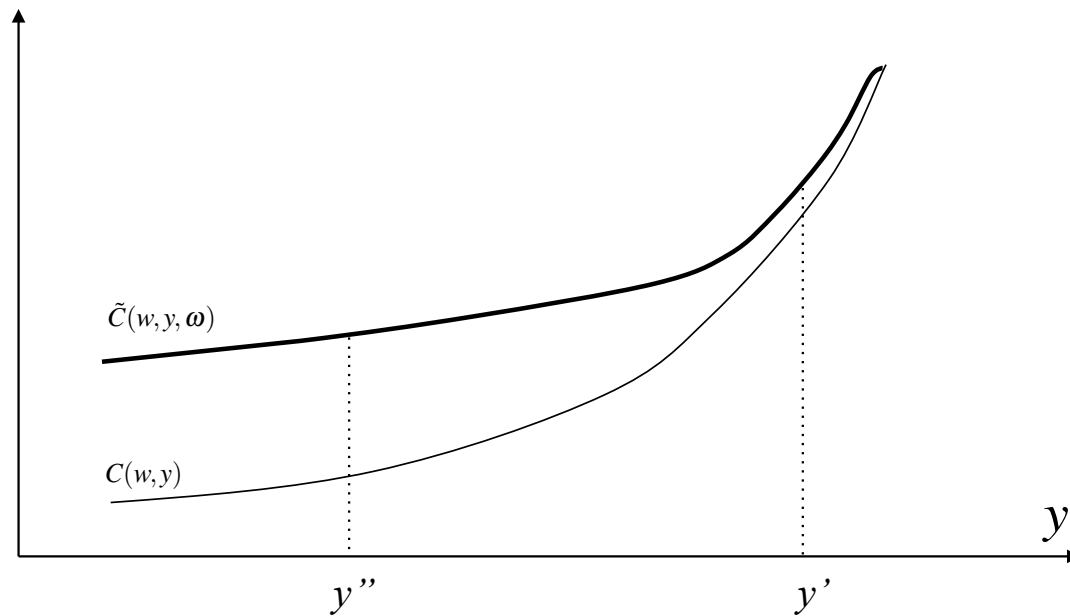
	Deterministic	Stochastic
Parametric	COLS	SFA
Non-parametric	DEA	RDEA StonED
Normative	Reference models	Simulation models

R1 corollaries

- › The cost norm is based on **evidence**
- › Data is filtered for **infeasible outliers**
- › The cost norm has **no extrapolations**
- › Data correspond to **estimation type (det/stoch)**



R2. Neutrality



Neutrality: there is no bias in $\tilde{C}(w, y, \omega)$ across the set ω .

$$E[\tilde{C}(w, y | \omega) / C(w, y)] = K < \infty \text{ for all } w, y \text{ in } \omega.$$

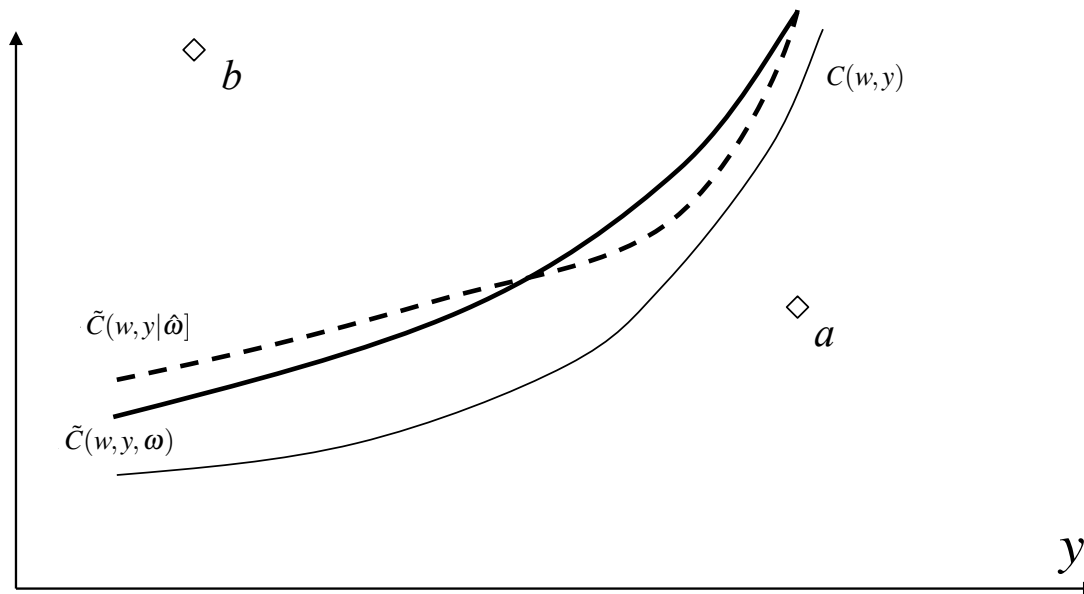
R2 Neutrality evaluations

	Deterministic	Stochastic
Parametric	COLS	SFA
Non-parametric	DEA	RDEA StonED
Normative	Reference models	Simulation models

R2 corollaries

- › The model specification has **no sample bias**
- › Norm error is **homoscedastic**
- › Efficiency distribution **independent of type**
- › Model specification **as comprehensive as possible**

R3. Robustness



Robustness: an estimate $\tilde{C}(w, y, \omega)$ is insensitive to irrelevant, flawed or inefficient data ε added to ω .

$$E[\tilde{C}(w, y|\omega) - \tilde{C}(w, y|\hat{\omega})] = 0 \text{ for all } w, y \text{ in } \omega, \hat{\omega} = \omega + \varepsilon.$$

R3 Robustness evaluations

	Deterministic	Stochastic
Parametric	COLS	SFA
Non-parametric	DEA	RDEA StonED
Normative	Reference models	Simulation models

R3 corollaries

- › The cost norm is **invariant to inefficient noise**
- › Data is filtered for **outliers or mavericks**
- › Frontier units are **identified and verifiable**

R4. Repeatability

Repeatability: insofar the estimation of $\tilde{C}(w, y, \omega)$ would require the determination of technical constants, assumptions and numerical parameters, they are endogenous or set *a priori* based on established practice and science. The calculation can be repeated with exactly the same results by a different set of equally skilled analysts.

R4 Repeatability evaluations

	Deterministic	Stochastic
Parametric	COLS	SFA
Non-parametric	DEA	RDEA StonED
Normative	Reference models	Simulation models

R4 corollaries

- › The cost norm has **no ad hoc assumptions**
- › No numerical **estimation procedures**
- › Reference set **endogenously defined**

Université catholique
de Louvain

UCL

HETEROGENEITY

Principles

Activity decomposition

- › Match process by process
- › Careful cost decomposition

Cautiousness in method

- › Minimal extrapolation methods
- › Data quality procedures

Cautiousness in estimation

- › Cost adjustments whenever relevant

Validation procedure

- › National validation by NRA
- › Validation with average European costs
- › Qualitative technical validation with detailed data
- › Cross validation by ranges, patterns, cost impact

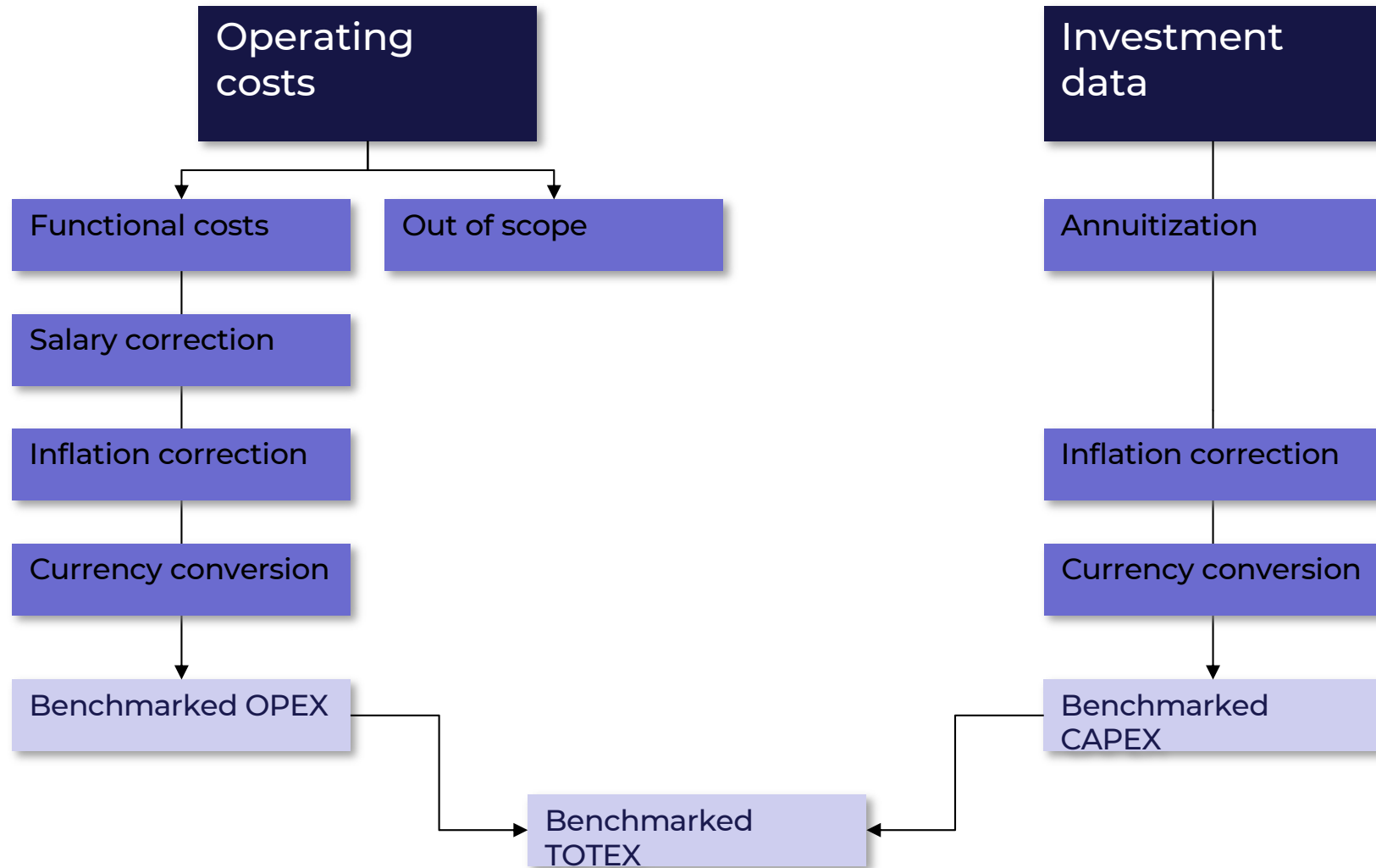
Process

Data validation

Structural comparability

- › Costs (functional decomposition)
- › Indirect cost allocation
- › Personnel costs
- › Time (inflation)
- › Assets (standard types, lives and functions)

Environmental comparability



Two approaches

Standard (average) cost

- › Specific claims for difference to average
- › Less initial data, fast average results
- › Burden of proof on TSO

Normative cost models

- › Using cost estimates from engineering
- › Higher initial cost, longer project
- › TSO input on open standards

Advanced environmental analysis TCB21

Condition	Parameters	Capex	Opex	Var
Land use cover	44 in 5 groups	X	X	yShare.area
Ruggedness (topography)	2	X		zRugged
Slope (topography)	4	X		zSlope
Subsoil (gravel, topsoil)	4	X		zGravel_T/S
Subsoil (gravel, subsoil)	4	X		zGravel_S
Humidity	5	X	X	zHumidity
Wetness index	1	X	X	zWWPI
Subsoil (depth to rock)	4	X		zSoil.dr
Subsoil (volume of stone)	4	X		zSoil.vs
Density	2	X		zDensity_po/co
Climate (extreme temperature)	8	X	X	zTemp_min/max
Climate (precipitation)	4		X	zPrec
Climate (snowdays)	1		X	zSnow
Soil water regime (subsoil wetness)	5	X	X	zWR

Université catholique
de Louvain

UCL

PROCESS

Principles

Data quality is primary objective

Data information management

Clear and fair communication

Results well explained, documented, justified

Data collection

- › Data inventory
- › Data collection process
- › Planning issues
- › Auditing concerns

Data collection guides

Auditing guides

Data validation protocols

Cross validation

Data quality assurance

Preparation prior to benchmarking:

- › Data definitions
- › Data guides
- › Collection tools with error checking

Important with TSO/NRA reviews

During project:

- › Feedback for validation
- › External validation
- › Dissemination of interim calculations for TSO control

Université catholique
de Louvain

UCL

OUTLIER DETECTION

Lack of frontier robustness

Structural

- › Heterogeneous and incommensurate technologies in the reference set
- › Operation at widely different scale in the reference set
- › Variations in scope of operation across the sample

Behavioral

- › Collusion (strategic reallocation of costs and outputs)
- › Maverick reporting

Source and intention of model error

Table 2: Four possible source for lack of robustness in Frontier-based models in regulation

Source of uncertainty	→Intention Level	
	Low	High
Firm	Reporting errors	Maverick reporting
Regulator	Poor variable specification	Biased model design

Maverick reporting

A model-based regulation needs to pass a **credibility test** at its launch

It may be in the interest of firms to provoke a **stress-test** for the regulation by intentionally creating **unrealistic peers** through strategic reporting.

The misreporting is then revoked due to 'reporting errors', but only after the frontier has been established.

Outlier detection

1. Superefficiency test
2. Dominance test
3. Peeling test

Université catholique
de Louvain

UCL

CONCLUSIONS

Conclusion

Acceptability = method compliance

- › **Feasibility** (evidence based estimate)
- › **Neutrality** (unbiased estimate)
- › **Robustness** (to data errors)
- › **Repeatability** (endogeneity of parameters)

Model objective = Adequacy

- › The cost norm is based on **structural comparability**
- › The resulting rulings should be **robust to judicial recourse**

Real practice more open for new techniques (hybrid) than most academic work

Main challenges

- › **Method integration** through a consistent **approach**
- › Addressing **heterogeneity**



Prof. Per J. AGRELL
CORE
UC Louvain
B-1348 Louvain-la-Neuve, BELGIUM
per.agrell@uclouvain.be

EXAMPLE: GAS MODELS

Early beginnings in electricity

Australia, BIE (1994)

New Zealand, IPART (1999)

ECOM (2002)

ECOM+ (2003, 2005)

E3GRID (2009)

E3GRID (2013)

TCB18 (2018)

TCB22 (2021)

Slow start in gas benchmarking

Carrington, Coelli, Groom (2002)

- › International benchmarking for gas distribution, NZ

Coelli, Estache, Perelman, Trujillo (2003)

- › World Bank book for regulators

Jamasb and Pollitt (2006)

- › Gas transmission
- › US data (FERC)
- › European data

Input orientation

Assumes that the major part of costs is incurred in capacity and infrastructure provision

The grid, an input, is seen as one of the outputs put at the disposal of the market

Output $y = f(\text{grid}, \text{context})$

Output is now

- › Partially endogenous
- › Linked to assets and capital investments
- › Technology dependent

Input orientation

Advantages

- › No utilization focus, promotes investments
- › Easy fit to cost functions (OPEX dependency)

Disadvantages for TSO applications

- › Gold plating or poor investment plans?
- › Risk for strategic choice of asset types?
- › Not linked to 'effectiveness'

Output orientation

Advantages

- › The configuration of the grid endogenous
- › Priority to relevant investments for use

Disadvantages for TSO applications

- › Part of output is forecast (load growth)
- › Utilization is strongly influenced of commodity prices and business cycle (post 2009 ...)
- › Capacity provision is not clearly identified, still a value

Hybrid orientation

Combination of input (grid provision) and output orientation (variable services, volumes, quality and flow)

Tradeoff between different services

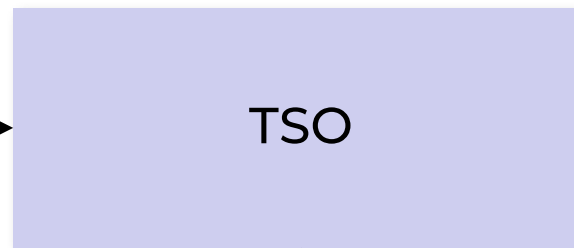
- › Capacity provision by asset provision
- › Transport work (flow and distance)
- › Customer service (connections, value-added services)

TSO benchmarking framework

1. Data standardization

X Inputs

Totex = Opex + Capex



Z Environment

2. Output modelling

Y Outputs

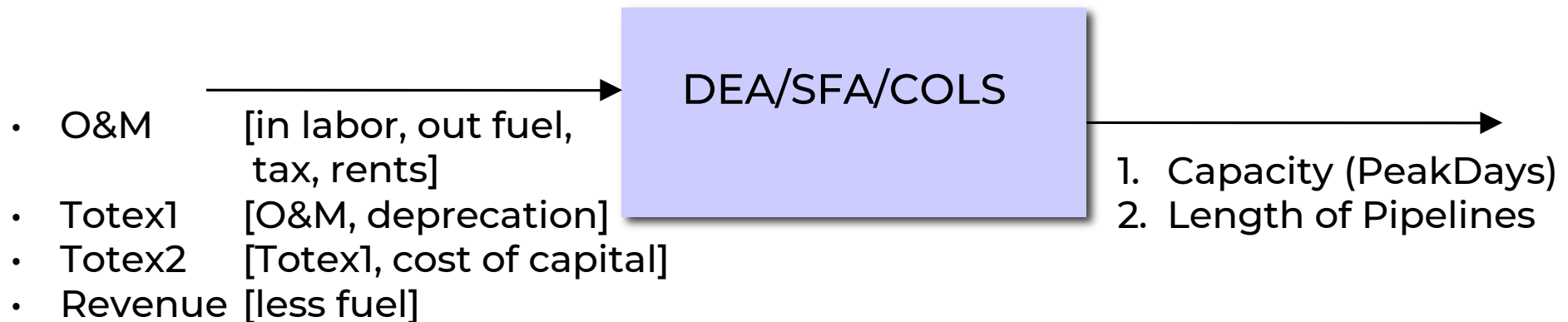
Transport work
Capacity provision
Service provision

3. Structural (environmental) modelling

Proxies for

- Geography, climate, soil type,
- Complexity, density
- ...

Jamasb & Pollitt, 2006

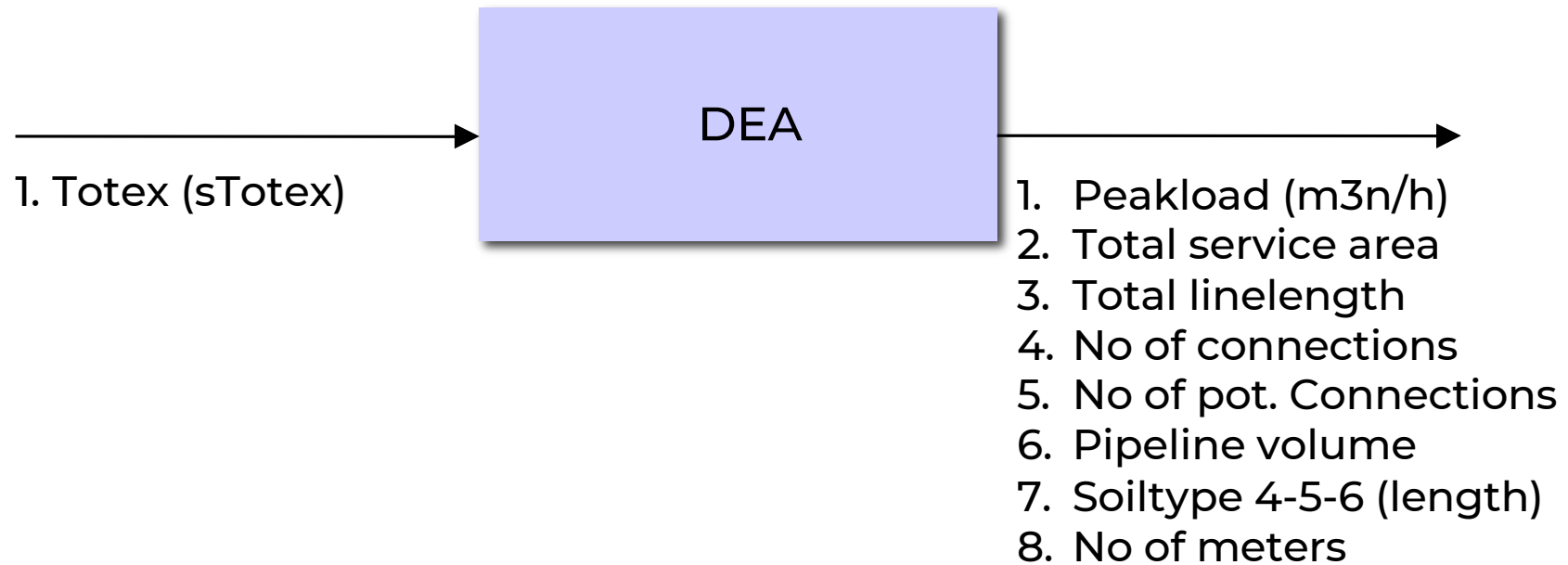


Data	Source:	CEER + FERC
	Country:	BE, DK, HU, NL, USA (43)
	Years:	1996-2005 (unbalanced)
	Observations:	47 TSO before data cleaning

Sumicsid, BNetzA - 2010



Data	Source:	Bundesnetzagentur (BNetzA)
	Country:	DE
	Year:	2010
	Observations:	10 RTO (1 year)



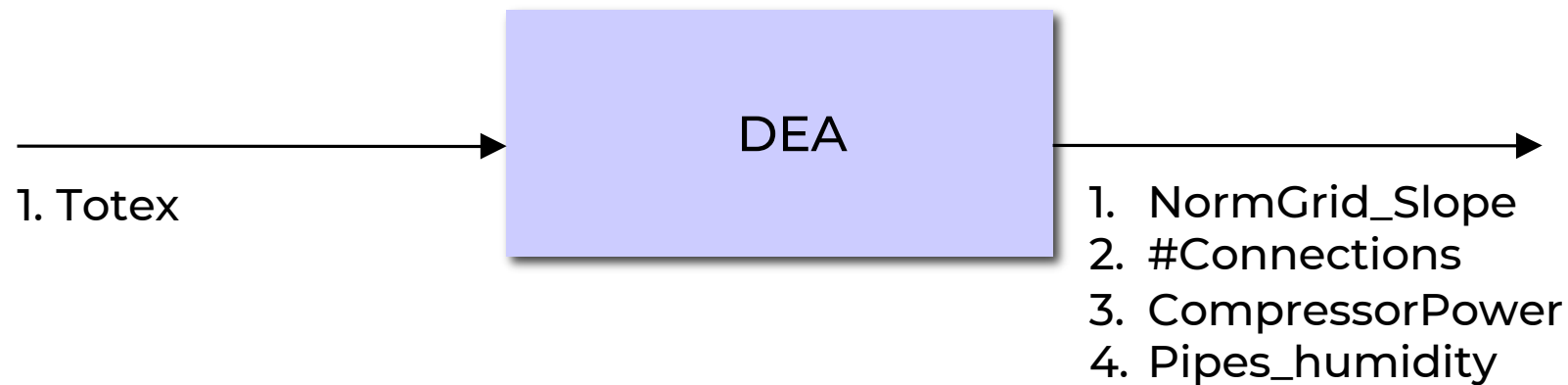
Data	Source:	Bundesnetzagentur (BNetzA)
	Country:	Germany
	Year:	2012
	Observations:	200 DSO before data cleaning

E2GAS, Sumicsid, CEER - 2016



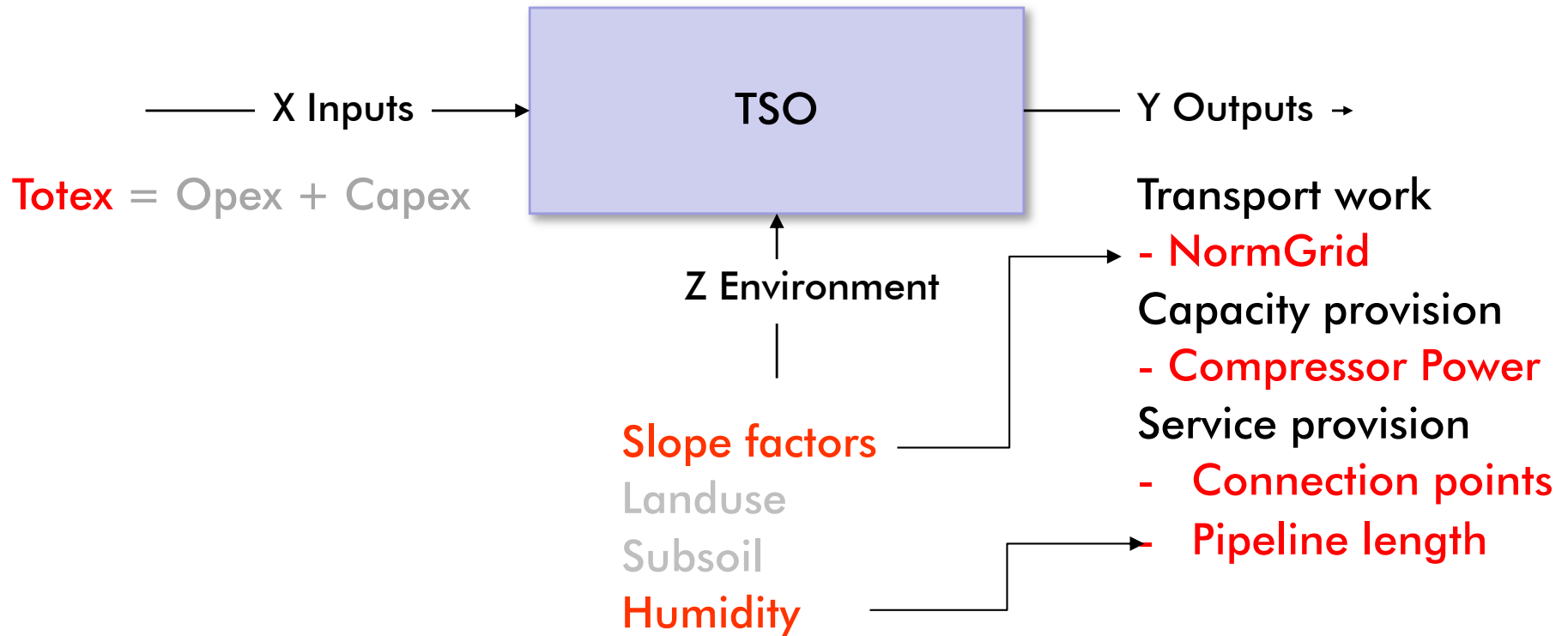
Data	Source:	CEER
	Country:	BE, DE, ES,FI, HR, NL, PT, UK
	Year:	2010, 2014
	Observations:	21 TSO (2 years)

TCB18, Sumicsid, CEER - 2018

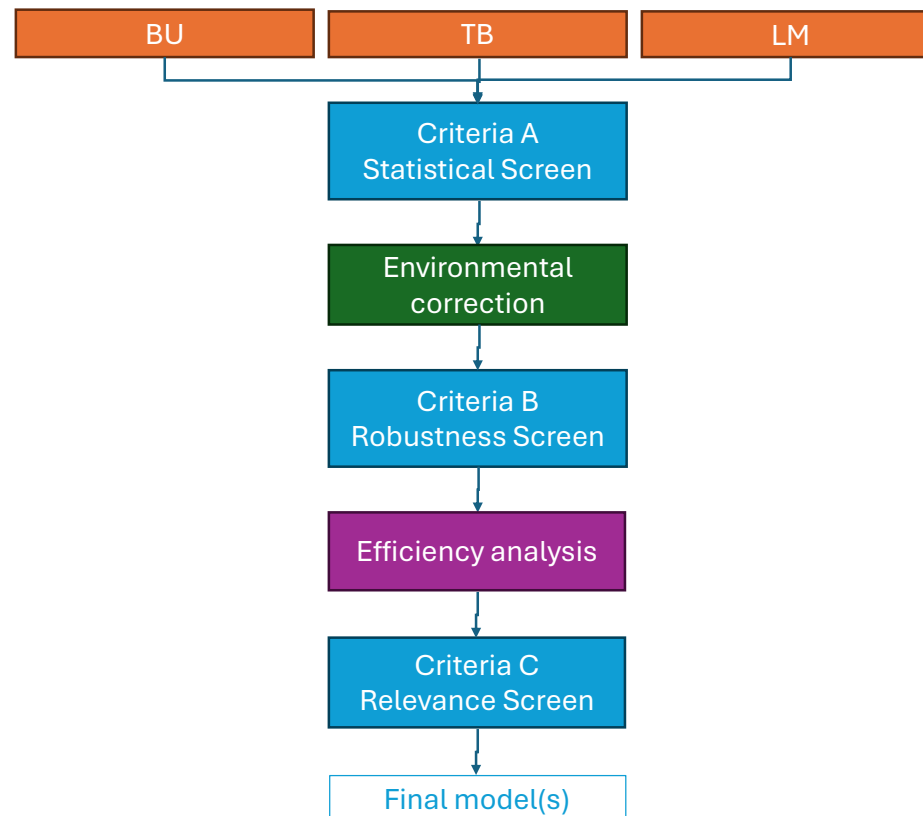


Data	Source:	CEER
	Country:	BE, DE, EE, ES, FI, GR, LT, LV, NL, PT, SI
	Year:	2017 (2013-2017)
	Observations:	29 TSO (2017, 2015 for DE), total 70

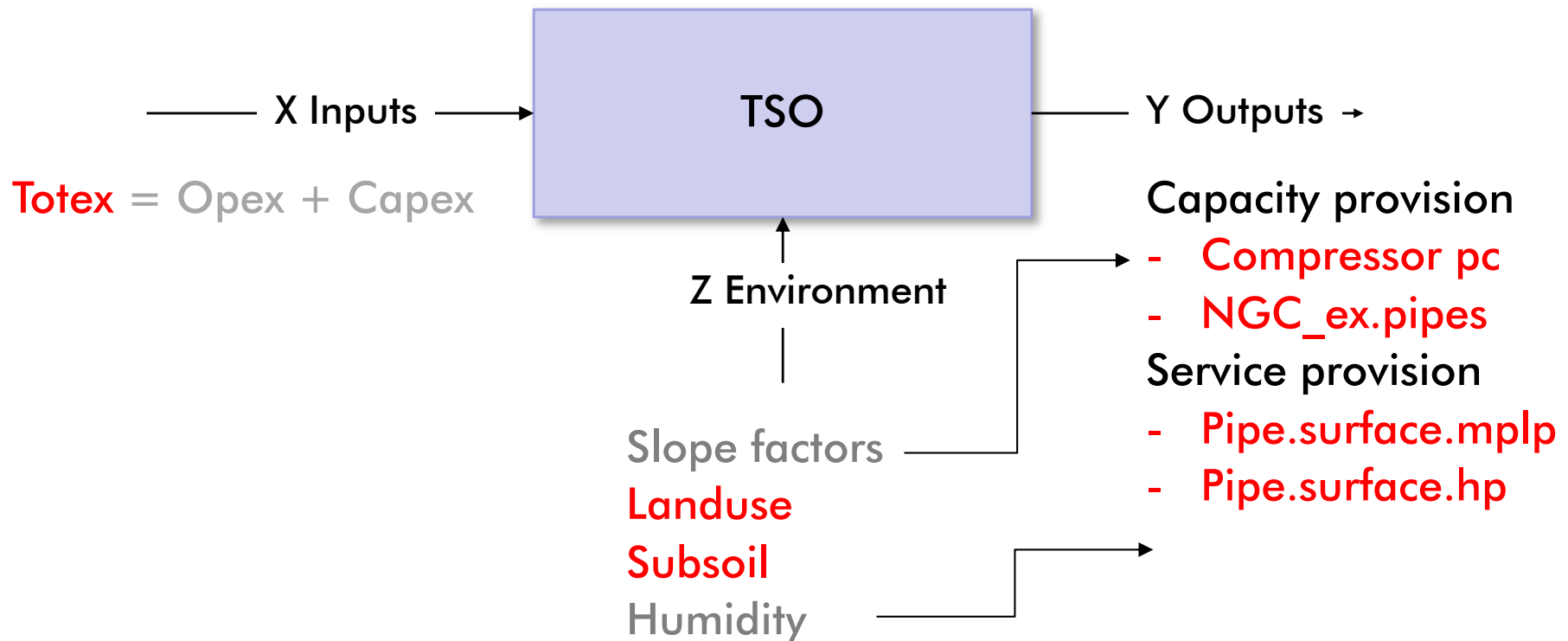
TCB18 Model (2018)



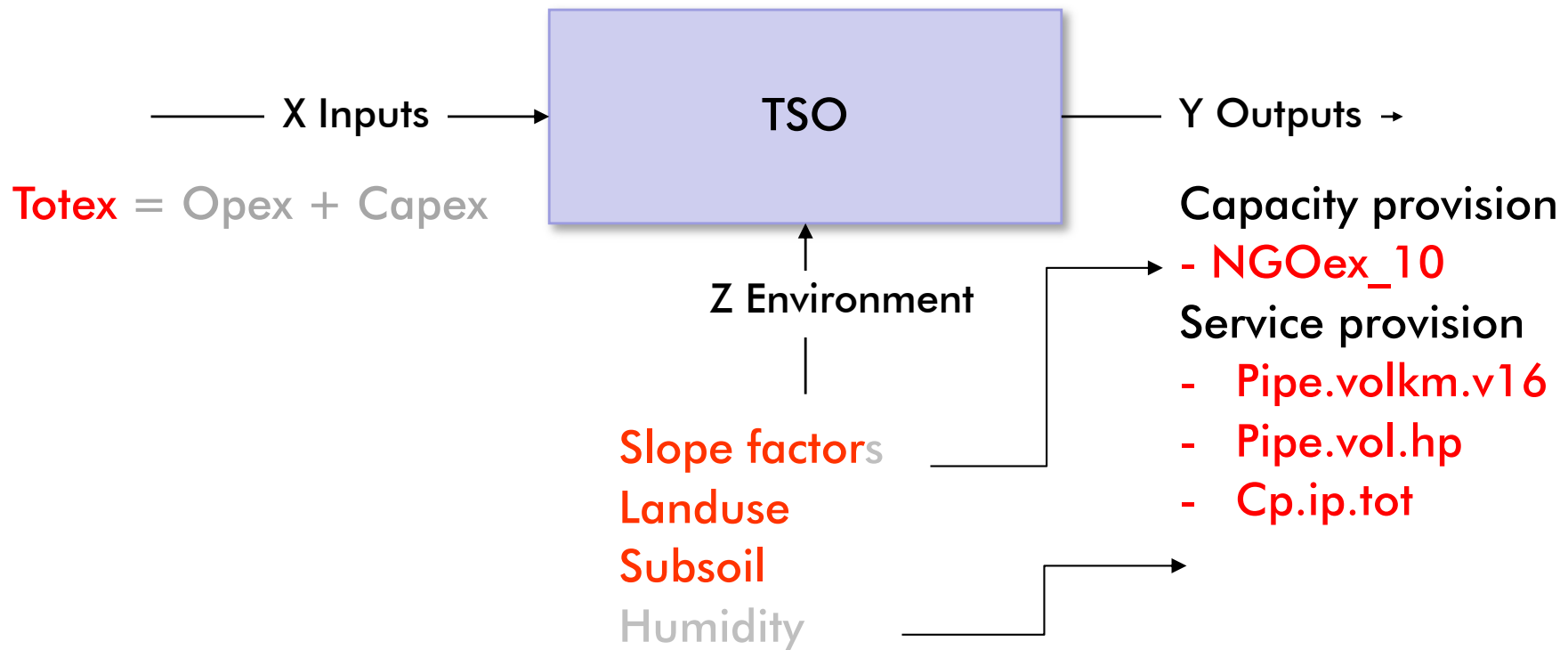
TCB21 Process



TCB21 Model G1445.V3 (2023)

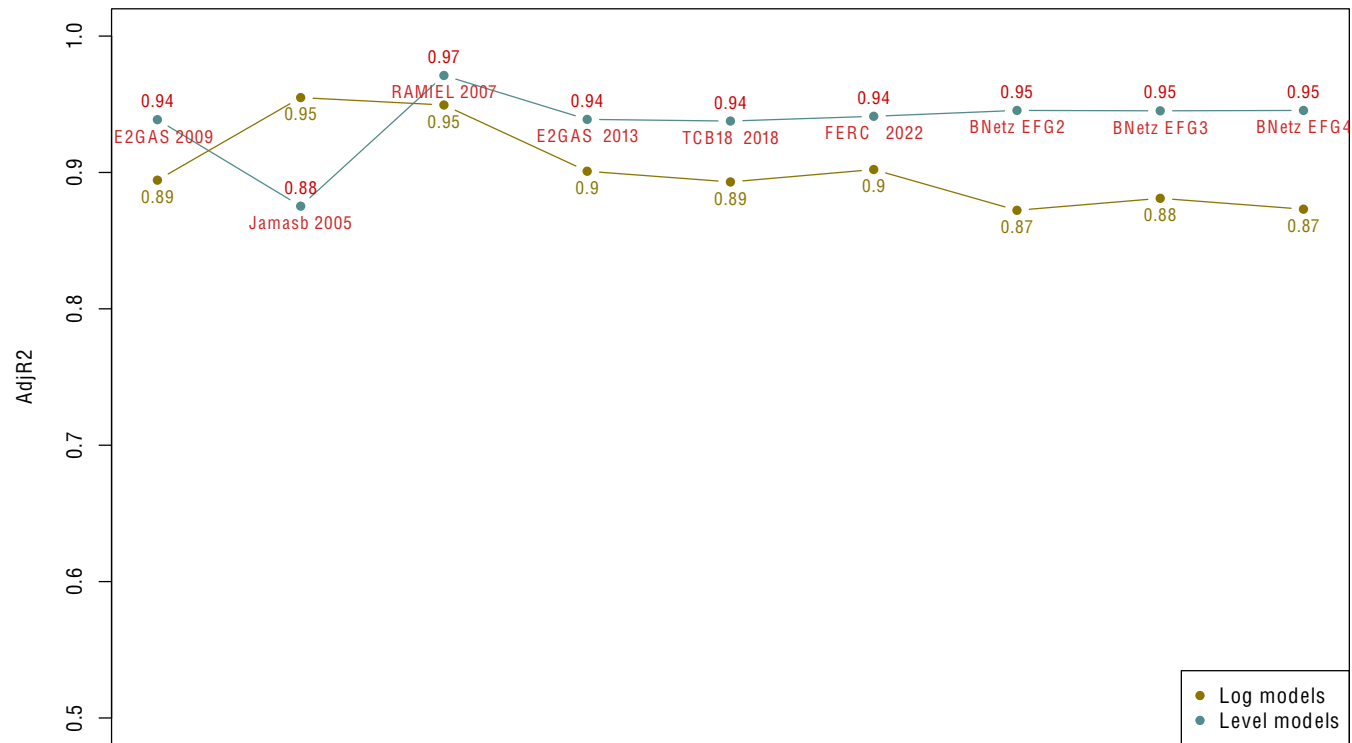


TCB21 Model G4450.V3 (2023)



Cross validation TCB21

TSOG legacy models / TCB21 data



projects

REFERENCES

References

- Agrell, P. J., & Bogetoft, P. (2007). Development of benchmarking models for German electricity and gas distribution. Consultation report, Bundesnetzagentur, Bonn, Germany, SUMICSID.
- Agrell, P. J., & Bogetoft, P. (2010). Harmonizing the Nordic regulation of electricity distribution. In *Energy, Natural Resources and Environmental Economics* (pp. 293-316). Springer Berlin Heidelberg.
- Agrell, P. J., & Bogetoft, P. (2018). Theory, Techniques, and Applications of Regulatory Benchmarking and Productivity Analysis. In *The Oxford Handbook of Productivity Analysis* (pp. 523-555). Oxford University Press.
- Agrell, P. J., Bogetoft, P., & Grammeltvedt, T. E. (2015). The efficiency of the regulation for horizontal mergers among electricity distribution operators in Norway. In proceedings from European Energy Market (EEM), 2015 12th International Conference on the (pp.1-5). IEEE.
- Agrell, P. J., Bogetoft, P., & Mikkers, M. (2013). Smart-grid investments, regulation and organization. *Energy Policy*, 52, 656-666.
- Agrell, P. J., P. Bogetoft & J. Tind (2002). Incentive plans for Productive Efficiency, Innovation and Learning. *International Journal of Production Economics*, 78, pp. 1- 11.
- Agrell, P. J, P. Bogetoft, & J. Tind (2005). DEA and Dynamic Yardstick Competition in Scandinavian Electricity Distribution, *Journal of Productivity Analysis*, 23, 173–201, 2005.
- Agrell, P. J., & Brea-Solís, H. (2017). Capturing heterogeneity in electricity distribution operations: A critical review of latent class modelling. *Energy Policy*, 104, 361-372.
- Agrell, P. J., & Niknazar (2014) Structural and Behavioral Robustness in Applied Best-practice Regulation. *Socio-Economic Planning Sciences*, 48(1), pp. 89-103.
- Agrell, P. J., & Grifell-Tatjé, E. (2016). A dynamic model for firm-response to non-credible incentive regulation regimes. *Energy Policy*, 90, 287-299.
- Agrell, P. J. & Lorenz, M. (2010) Calling the Cards: Fears, Threats and Delegation in Infrastructure Regulation Games. CRESSE Conference 2010 Advances in the Analysis of Competition Policy and Regulation, Chania, Greece. www.uclouvain.be
- Agrell, P.J. & Teusch, J. (2017) Strategic behaviour under frontier regulation. CORE Discussion paper, Université catholique de Louvain, forthcoming.
- Bogetoft, P. (1994), Incentive Efficient Production Frontiers: An Agency Perspective on DEA, *Management Science*, 40, pp.959-968.
- Bogetoft, P. (1995), Incentives and Productivity Measurements, *International Journal of Production Economics*, 39, pp. 67-81.
- Bogetoft, P. (2000), DEA and Activity Planning under Asymmetric Information, *Journal of Productivity Analysis*, 13, pp. 7-48.



Incentive regulation

Coelli T.J and Lawrence D. (Eds) (2006) Performance Measurement and Regulation of Network Utilities, Edward Elgar Publishing: Cheltenham.

Vogelsang, I. (2002). Incentive regulation and competition in public utility markets: A 20-year perspective. *Journal of Regulatory Economics*, 22(1), 5-27.

Weyman-Jones, T., J. Bouchina, C. Godinho, C. Feteira Inacio, P. Martins and J. Latore (2006) Efficiency Analysis for Incentive Regulation, in T.J Coelli and D. Lawrence (ed.), *Performance Measurement and Regulation of Network Utilities*, Edward Elgar Publishing: Cheltenham, 329-363.

Bibliography incomplete – see references in papers when indicated.

International benchmarking

- Carrington, R., Coelli, T., & Groom, E. (2002). International benchmarking for monopoly price regulation: The case of Australian gas distribution. *Journal of Regulatory Economics*, 21(2), 191-216.
- Coelli, T., Estache, A., Perelman, S., & Trujillo, L. (2003). *A primer on efficiency measurement for utilities and transport regulators*. Washington, DC: world bank.
- Haney, A. B., & Pollitt, M. G. (2009). Efficiency analysis of energy networks: An international survey of regulators. *Energy policy*, 37(12), 5814-5830.
- Haney, A. B., & Pollitt, M. G. (2011). Exploring the determinants of “best practice” benchmarking in electricity network regulation. *Energy Policy*, 39(12), 7739-7746.
- Haney, A. B., & Pollitt, M. G. (2013). International benchmarking of electricity transmission by regulators: A contrast between theory and practice?. *Energy Policy*, 62, 267-281.
- Jamasb, T., & Pollitt, M. (2000). Benchmarking and regulation: international electricity experience. *Utilities Policy*, 9(3), 107-130.
- Jamasb, T., & Pollitt, M. G. (2004). *Benchmarking and Regulation of Electricity Transmission and Distribution Utilities: Lessons from International Experience*. <https://doi.org/10.17863/CAM.5398>

Bibliography incomplete – see references in papers when indicated.

Multi-methods

Kuosmanen, T., Saastamoinen, A., & Sipiläinen, T. (2013). What is the best practice for benchmark regulation of electricity distribution? Comparison of DEA, SFA and StoNED methods. *Energy Policy*, 61, 740-750.

Coelli T.J and Lawrence D. (Eds) (2006) *Performance Measurement and Regulation of Network Utilities*, Edward Elgar Publishing: Cheltenham.

Coelli, T., Estache, A., Perelman, S., & Trujillo, L. (2003). A primer on efficiency measurement for utilities and transport regulators. Washington, DC: world bank.

Hjalmarsson, L., Veiderpass, A. (1992). Productivity in Swedish electricity retail distribution. *Scand. J. Econ.* 94, 193–205.

Jamasb, T., & Pollitt, M. G. (2004). Benchmarking and Regulation of Electricity Transmission and Distribution Utilities: Lessons from International Experience. <https://doi.org/10.17863/CAM.5398>

Waddams-Price, C. & Weyman-Jones, T. (1996). Malmquist indices of productivity change in the UK gas industry before and after privatization. *Applied Economics*, 28(1), 29-39.

Bibliography incomplete – see references in papers when indicated.

Quality

Giannakis, D., Jamasb, T., & Pollitt, M. (2005). Benchmarking and incentive regulation of quality of service: an application to the UK electricity distribution networks. *Energy policy*, 33(17), 2256-2271.

Bjørndal, E., Bjørndal, M., & Fange, K. A. (2010). Benchmarking in regulation of electricity networks in Norway: An overview. *Energy, natural resources and environmental economics*, 317-342.

Bibliography incomplete – see references in papers when indicated.