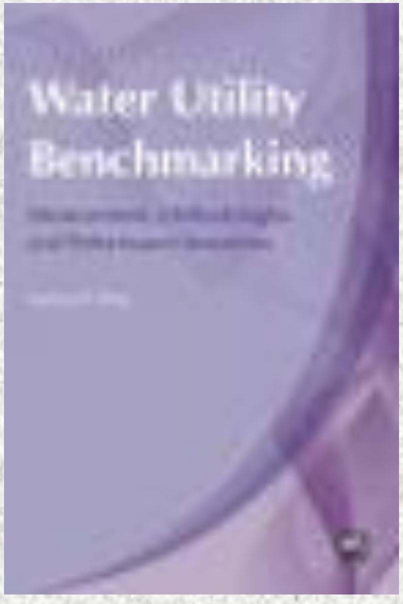
Conducting Benchmarking Studies



Sanford Berg

Public Utility Research Center

University of Florida

www.purc.ufl.edu

Symbolic and Substantive Importance

- Water and energy have come to symbolize the huge gap between promise and performance.
- Globally, lives are lost each day when serious reforms are delayed and investments deferred.
- Domestically, huge "investment gaps" have been identified in various reports.
- The political economy of infrastructure is such that those who make tough decisions will not receive credit during their terms in office.

Importance of Making Comparisons

Performance comparisons are necessary but not sufficient for sound policy

Benchmarking represents an important tool for

- Documenting past performance,
- Establishing baselines for gauging improvements,
- Making comparisons across service providers and over time, and
- Designing incentives.

Objectives of the IWA book

- Bridge the gap between technical researchers and practitioners currently conducting studies for government agencies and water utilities.
- Encourage the application of more sophisticated quantitative tools to promote policies that improve company (and sector) performance.
- Provide rigorous tools that allow stakeholders to quantify utility progress towards meeting policy objectives, help water specialists identify high performing utilities (whose production processes might be adopted by others), and enable regulators to fine tune targets and incentives for utilities.

Definition of Benchmarking

Benchmarking is an ongoing, systematic process to search for and introduce international best practice into your own organization, conducted in such a way that all parts of your organization understand and achieve their full potential. The search may be of products, services, business practices and processes, of competitors or those organizations recognized as leaders, in the industry or in the specific processes that you have chosen.

Denis Lawrence (Tasman Asia Institute)

Focus on Performance Metrics

This World Bank funded study shows how analysts can

- ***measure*** water utility ***operations*** (costs, physical inputs and outputs)
- **to *perform company comparisons***
- in the **context of *infrastructure reform***.

The focus is on Performance Scores and Rankings based on quantitative production and cost studies.

Elements of Metric Benchmarking

- Collect Data and Establish baselines (past performance) as a starting point.
- Prepare quantitative comparisons using cross-sectional and time-series analysis.
- Identify Relative Performance (ensure that firms face comparable conditions).
- Devise incentives so cost-savings are ultimately passed on to customers.
- Promote managerial strategies to achieve best practice.

Data Collection: Global and National

- Water/sewerage system operations,
- Network capacity,
- Financial flows, and Outputs.

Consistent data are essential for good management and for public policy oversight.

For Electricity: See Jamasb & Pollitt (2001) "Benchmarking and Regulation," *Utilities Policy* 9 (3).

Comparable data facilitating cross-country comparisons: *Water & Sanitation International Benchmarking Network* (IBNET, funded by the UK Department for International Development and the World Bank).

Can an Index Capture Complexity?

A single index of utility performance will be neither comprehensive nor fully diagnostic.

- Physician can have information on a patient’s temperature, pulse, height and weight.
- Patient is in trouble: dangerous fever and/or is significantly overweight.
- Blood tests provide more detailed information
- Diagnosing and treating mental health issues would require other diagnostics and treatments . . . Still, temperature and weight provide useful information.

Five Benchmarking Methodologies

- Core Indicators and a Summary or Overall Performance Indicator (partial metric method),
- Performance Scores based on Production or Cost Estimates ("total" methods),
- Performance Relative to a Model Company (engineering approach),
- Process Benchmarking (involving detailed analysis of operating characteristics), and
- Customer Survey Benchmarking (identifying customer perceptions).

KPIs and Overall Performance Indicators

- *Specific Key Indices*, such as water delivered per worker, quality of service (continuity, water quality, complaints), unaccounted for water, coverage, and key financial data (operating expenses relative to total revenues, collections).
  - partial measures provide the simplest way to perform comparisons: trends direct attention to potential problem areas

- *Overall Performance Indicator (OPI)* combines the specific core indices into a summary index
  - Peru: **9 KPIs**: summed by SUNASS (3 quality, 2 coverage, 3 management, 1 finance (Cost/Revenues)
  - Canada: **61 KPIs**

Performance Scores Based on Production or Cost Estimates

Rankings can be based on the analysis of production patterns and/or cost structures.

Production function studies (requiring data on inputs and outputs) show how inputs affect utility outputs (such as volume of water delivered, number of customers, and service quality). Utilities that produce far less output than other utilities (who are using the same input levels) are deemed to be relatively inefficient.

Cost functions show how outputs, inputs and input prices affect costs; such models have heavy data requirements. Excessively high costs would trigger more in-depth studies to determine the source of poor performance.

Use of Performance Scores

Identify the best performers and the weakest performers in a group of utilities:

- Technical/engineering efficiency,
- Cost efficiency,
- Production efficiency,
- Scale efficiency,
- Economies of scope (water and wastewater) or transmission and distribution)
- Efficiency change

Engineering/Model Company

- Requires the development of an optimized economic and engineering model
- Idealized benchmark specific to each utility—incorporating the topology, customer demand patterns, and density of the service territory.
- “Artificial” firm has optimized its network design and minimized its operating costs
- Production relationships can be obscured through a set of assumed coefficients used in the optimization process.
- Chile and Argentina used this approach to establish infrastructure performance targets.
- US telecom interconnection pricing and battles of “models”

Process Benchmarking

- Focuses on individual production processes
- Detailed examination of facilities and their operations
  - Identifies stages of the **production process** needing attention: pumping up, intake, transport, clarification and filtration, purification and treatment.
  - Studies of **distribution processes** (network design, pipeline construction and maintenance), **sales processes** (meter reading, data processing, billing, collections, and customer relations), and **general processes** (planning, staff recruitment and retention, and public relations).
- Provides a mechanism for identifying potential benchmarking partners, undertaking benchmarking visits, and implementing best practices

Mats Larsson, et al (2002). *Process Benchmarking in the Water Industry: Towards a Worldwide Approach*, IWA

Customer Survey Benchmarking

- Customer Complaints: one indicator
- SERVQUAL identifies five dimensions of service quality as perceived by customers:
  - **External characteristics** (tidy workplace, employee appearances),
  - **Reliability** (meeting deadlines, consistency in interactions),
  - **Responsiveness** (providing service promptly),
  - **Consideration** (personnel who are courteous, friendly, and helpful),
  - **Empathy** (giving individual care and attention).

Parasuraman et. al. (1985) *Journal of Marketing*

R. Parena (1999). *The IWSA Benchmarking Initiatives:*

Steps for Conducting Production or Cost Benchmarking Studies

1. **Identify objectives, select methodology, and gather data;**
2. **Screen and analyze data;**
3. **Utilize specific analytic techniques;**
4. **Conduct consistency/sensitivity tests;**
5. **Develop policy implications.**

Step 1. Identify Objectives, Select Methodology and Gather Data

- Decide issues to be addressed, time period to be analyzed, and types of comparisons
- Choices will reflect capabilities, initial understanding of data availability, and preliminary methodological choices.
- The objectives of any benchmarking study will depend on most important policy issues under consideration.
- Staff requirements can be substantial

Step 2. Screen and Analyze Data

- Screen data
  - timeframe,
  - sample size, and
  - statistical techniques.
- Check data quality
  - inconsistent definitions,
  - missing data or
  - extreme data values
- Analysis is an iterative process

Step 3. Utilize Specific Analytic Techniques

Based on data and objectives, choose technique

- Core Overall Performance Indicators
- Total Factor Productivity Indices
- Relative Performance using Data Envelopment Analysis (DEA)
- Relative Performance using Statistical Techniques

Utilize specialists for sophisticated models

“A Primer on Efficiency Measurement for Utilities and Transport Regulators”, by Coelli, Estache, Perelman, & Trujillo, World Bank Institute

Metric Methodologies: Four Categories

Core Overall Performance Indicators (OPIs combine partial indicators of operating or financial performance; these are summary indices),

Total Factor Productivity Indices (an index number approach using output per unit input—multiple inputs are taken into consideration to gauge efficiency levels and changes),

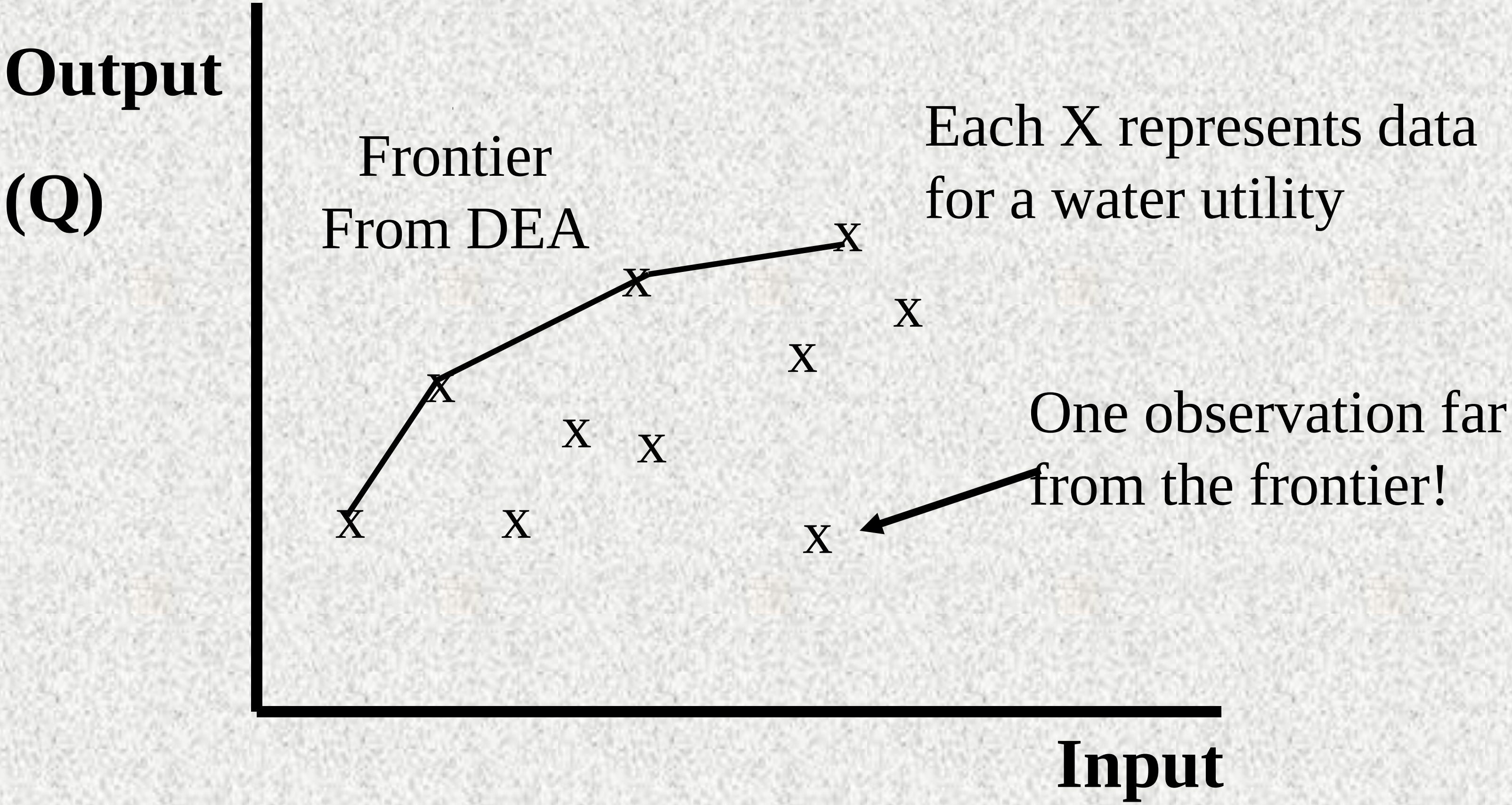
***Relative Performance using Data Envelopment Analysis** (a non-parametric technique that makes no assumptions about the functional form of production or cost functions),

***Relative Performance using Statistical Techniques** (parametric approaches that involve assumptions about functional relationships)

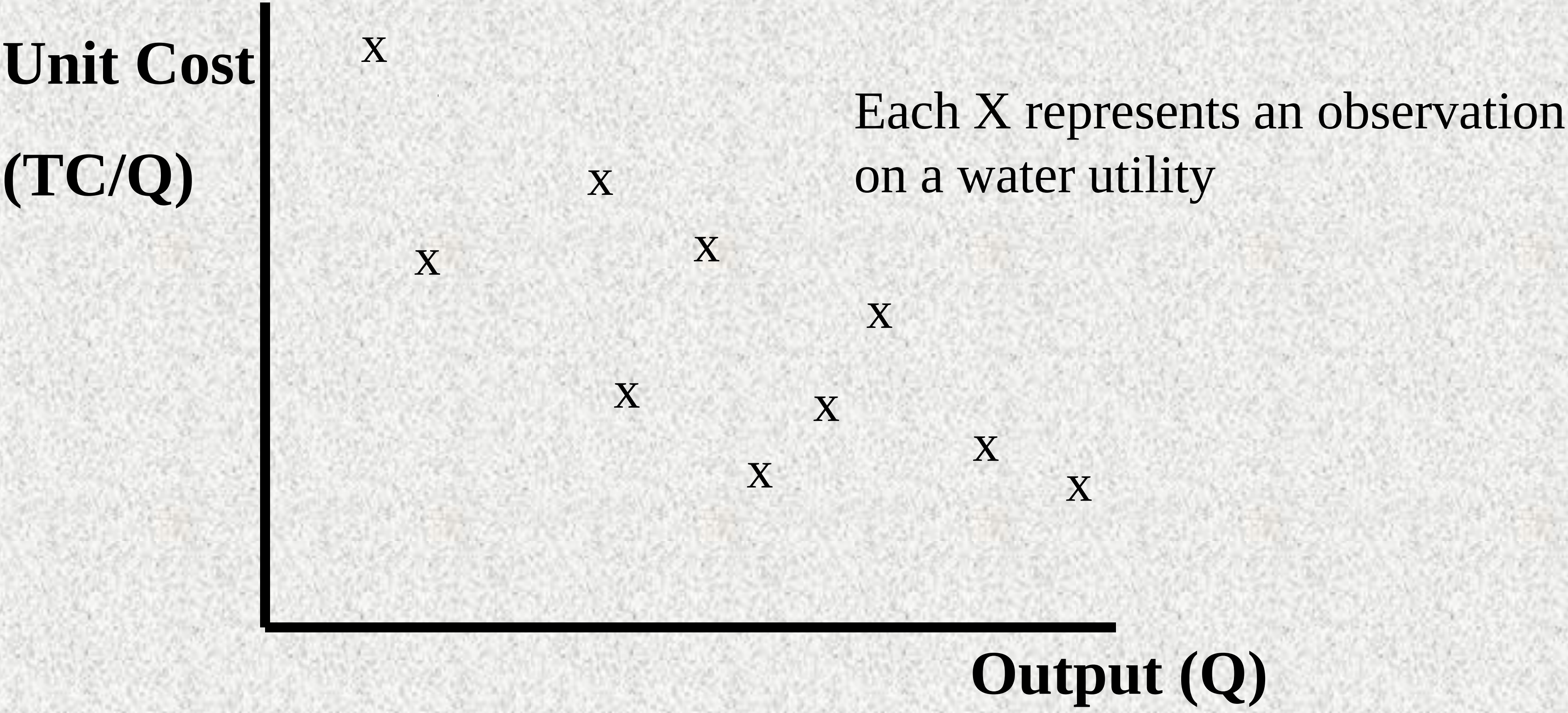
Data Envelopment Analysis (DEA)

- Linear programming technique applied to a selected set of variables to calculate an efficiency coefficient for each water utility.
- Adapted from the multi-input, multi-output production function,
  - Production Function: how much output can be produced with a given basket of inputs.
  - DEA benchmarks firms only against the best producers (so it is a non-parametric frontier analysis). Differs from OLS that bases comparisons with respect to an average producer.
- Applied to cost functions as well

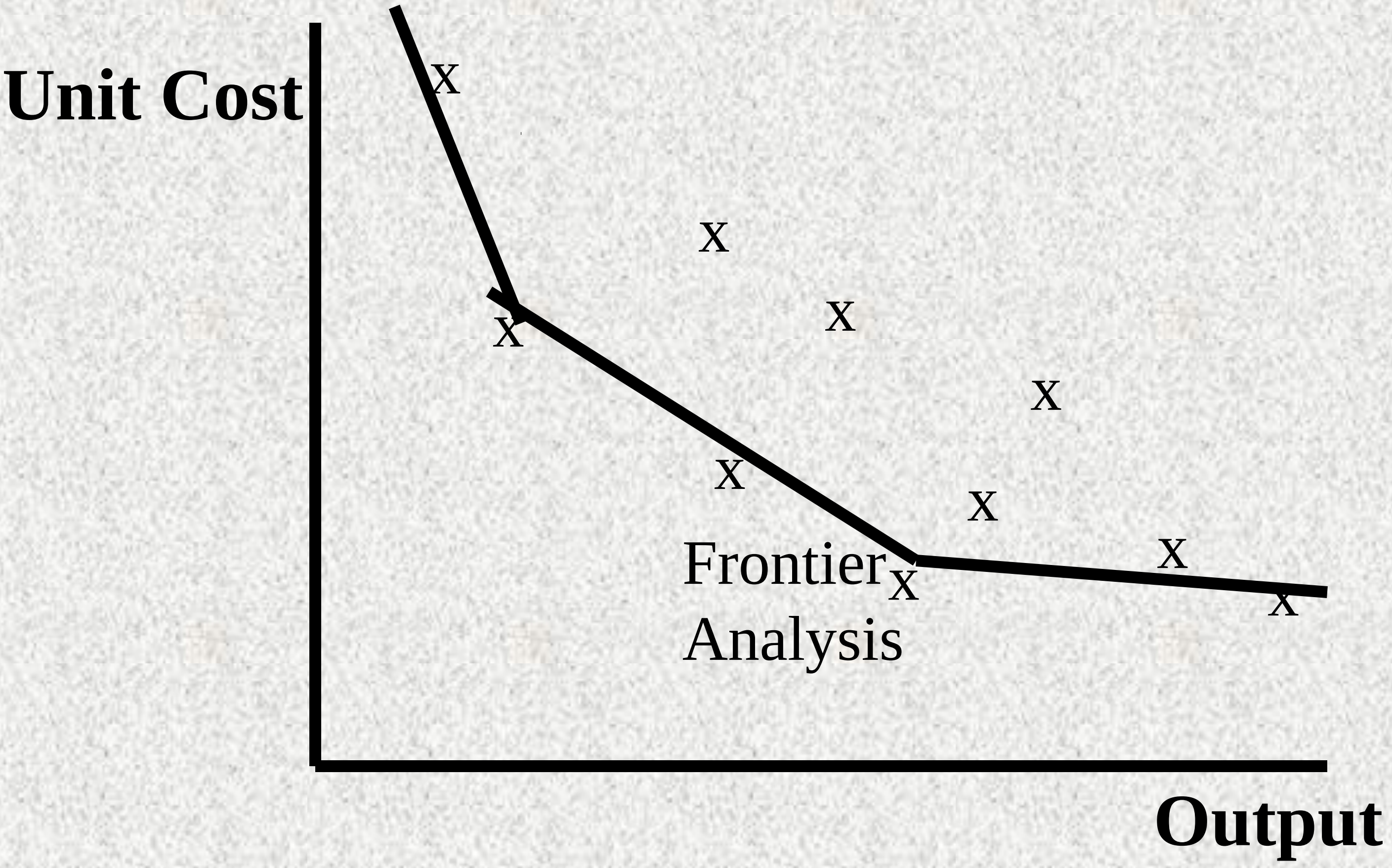
Production Observations



Unit Cost Observations



Frontier Analysis



Statistical Analysis

Identify relationship between firm performance and market conditions and characteristics of the production processes.

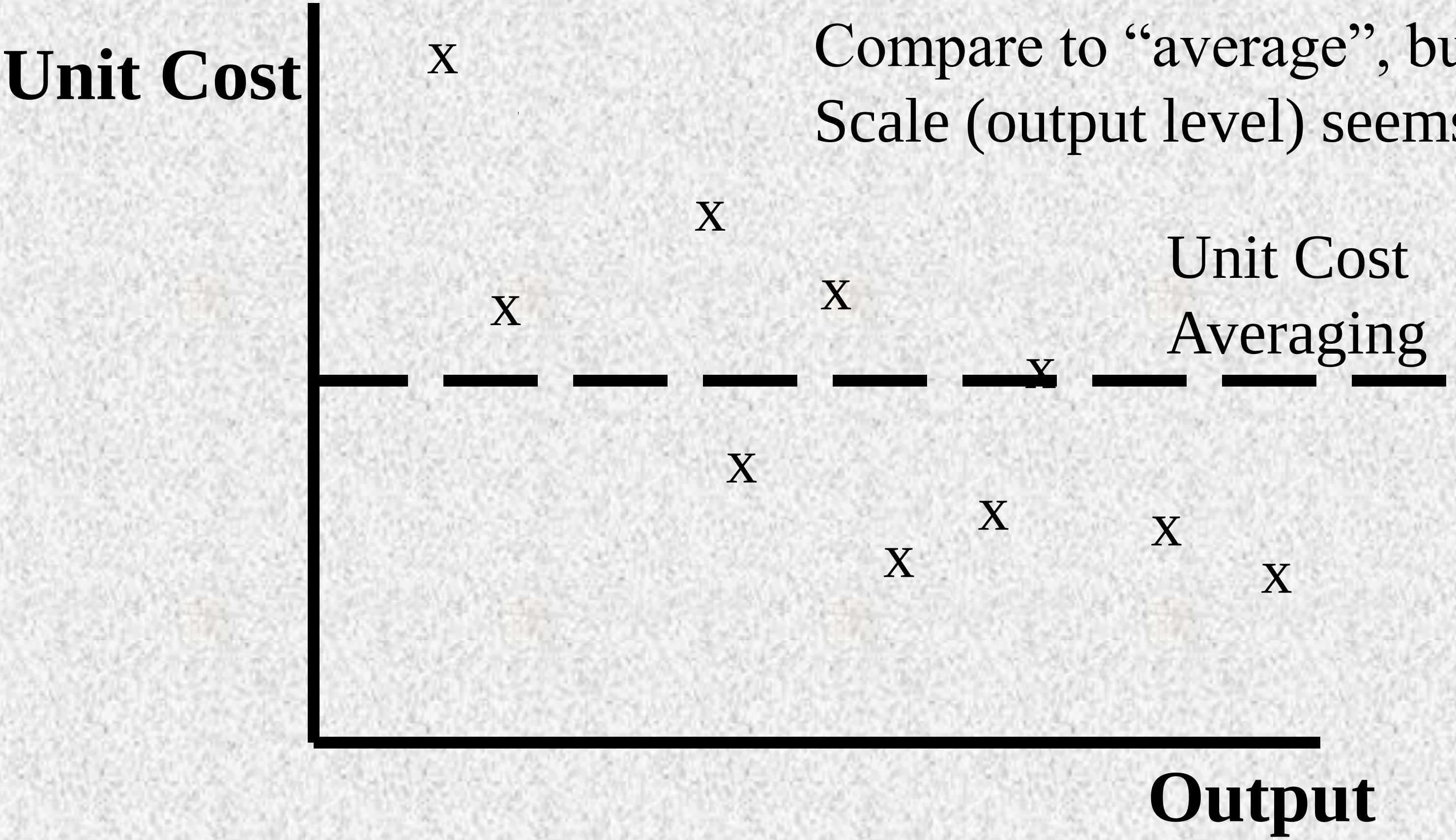
- Coefficients: the roles of multiple variables
- Calculated expected cost, given actual output.
- Predicted versus actual cost can provide a measure of relative performance.

Widely used econometric techniques include

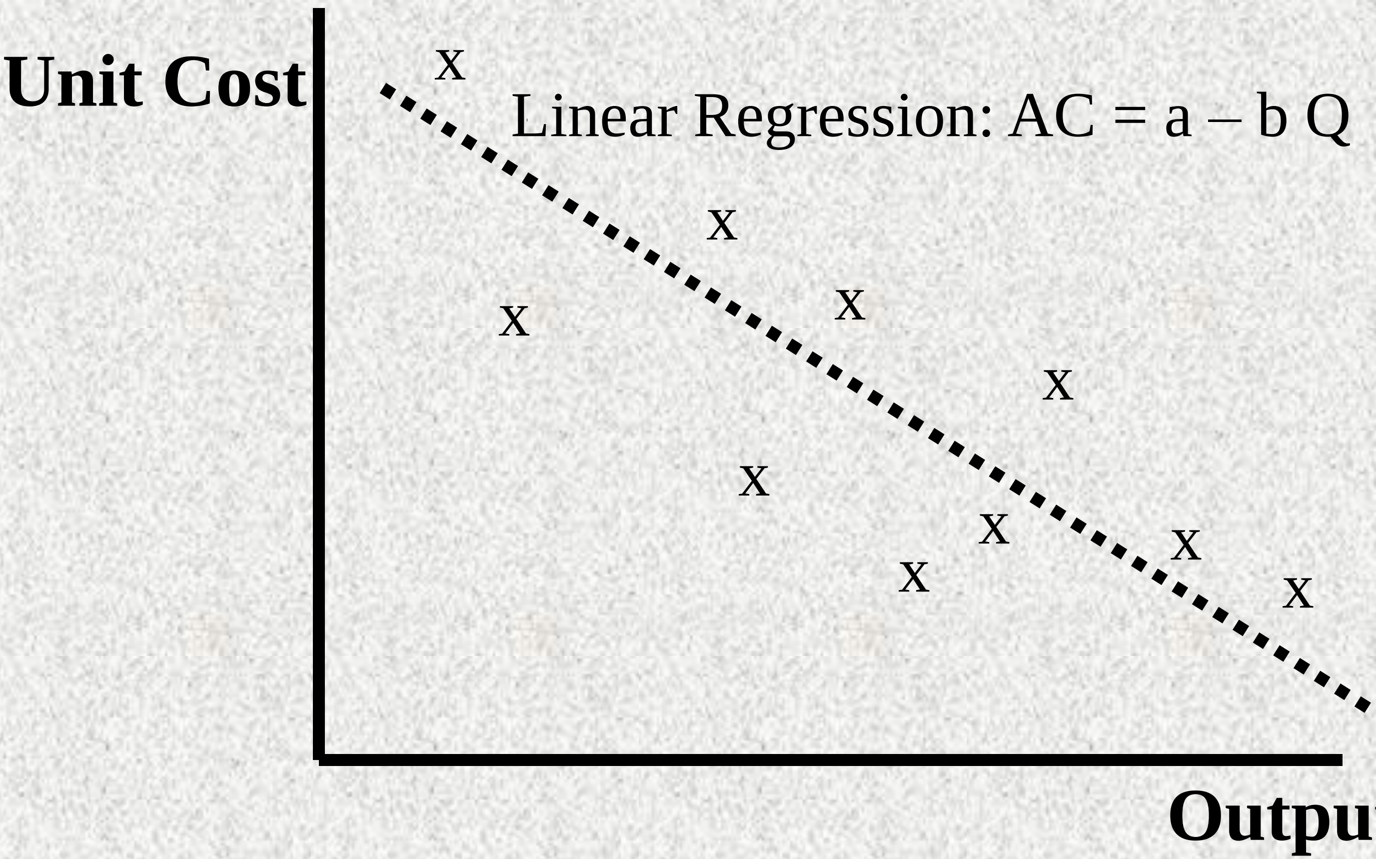
- *Ordinary Least Squares (OLS)*
- *Corrected OLS* (COLS)
- *Stochastic Frontier Analysis* (*SFA)*.

Unit Cost Averaging

Compare to “average”, but **here**
Scale (output level) seems to matter

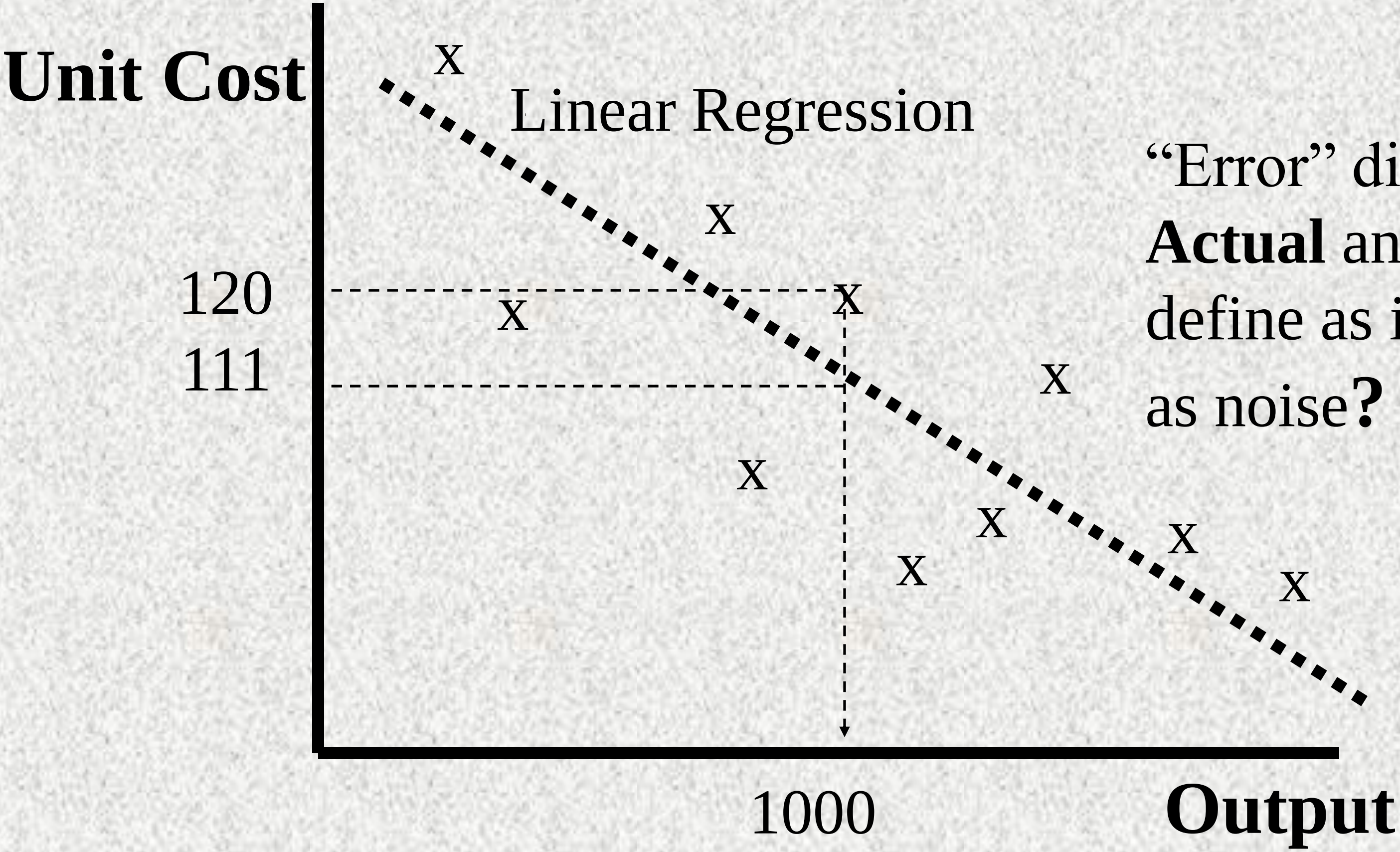


Linear Regression



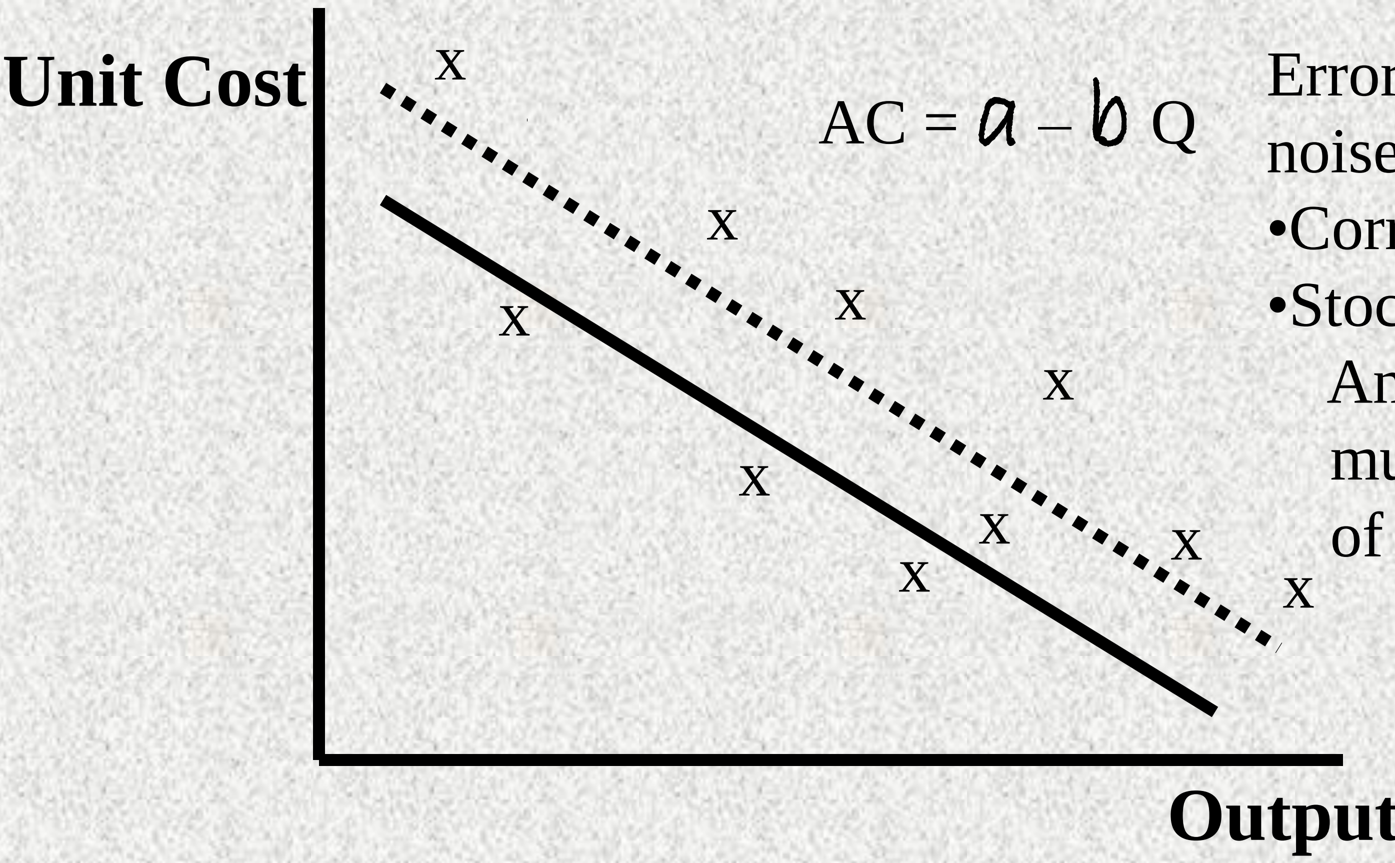
- Functional form
- Variables
- Error terms

Relative Inefficiency (OLS?)



"Error" difference between **Actual** and **Predicted** Cost: define as inefficiency or as noise**?**

Frontier Regression



Error terms (random noise?)

•Corrected OLS

•Stochastic Frontier Analysis (SFA, with multiple components of error term)

Model Specification

Functional forms:

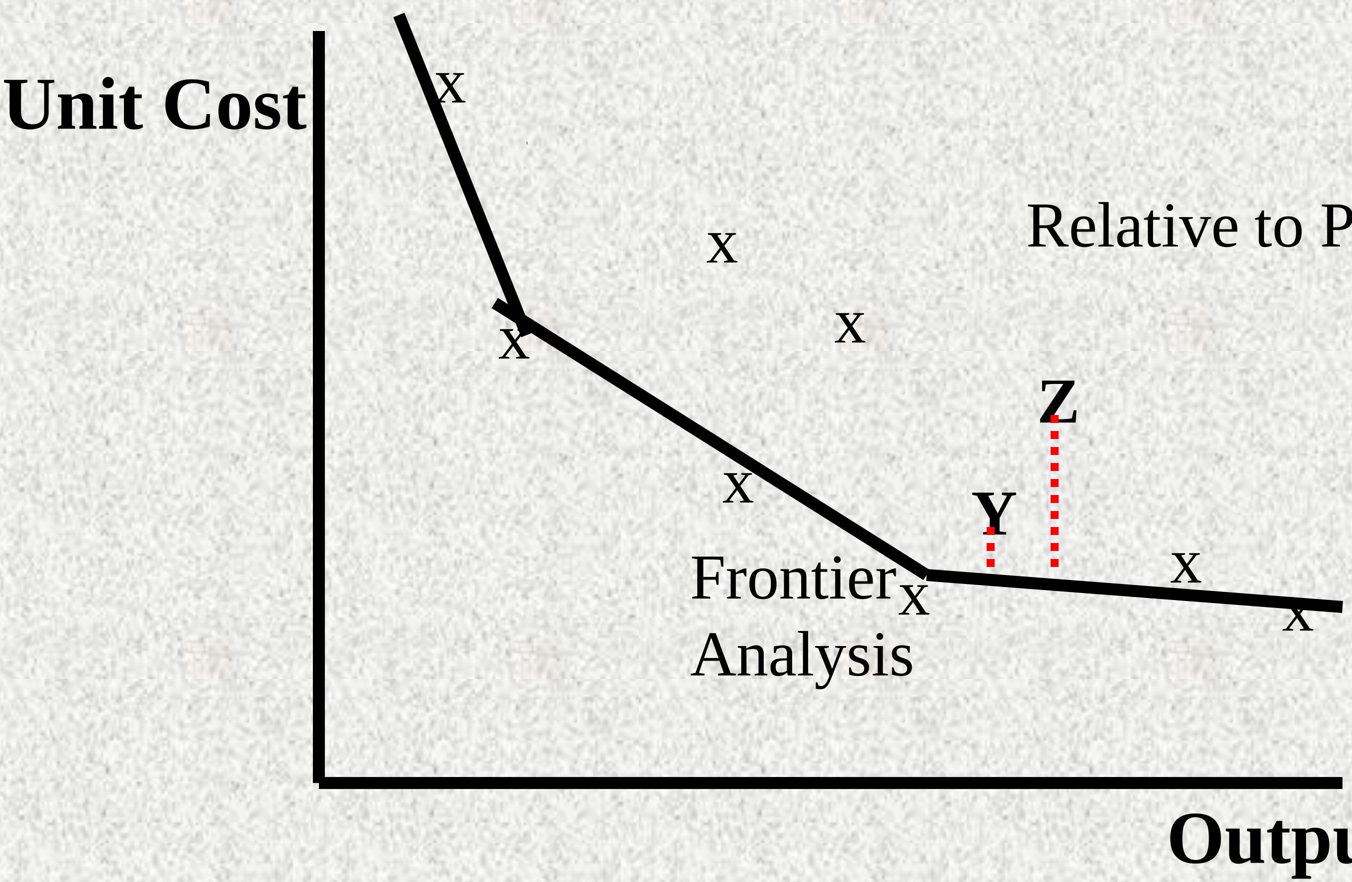
- Linear or non-linear
- Interpretation of error terms

Cross Section, Time Series, Panel?

Excluded Variables?

- Service Quality
- Customer Density? Age of network?
- Topography, Hydrology, Ownership
- Input prices (for Cost function)

Relative Inefficiency (frontier/non-linear)



Relative to Predicted Cost

Develop
- Scores
- Rankings

Cost Example: Ofwat Water Service Model (Y = Opex*)

Operating expenditures (less “exceptionals”, rates, third party services, abstraction charges, pumping costs)

$$\text{Ln } Y = 3.57 + 0.471 \text{ Ln } X_1 + 0.468 \text{ Ln } X_2 - 1.575 \text{ Ln } X_3$$

X1 = water delivered in Ml/day

X2 = length of main in km

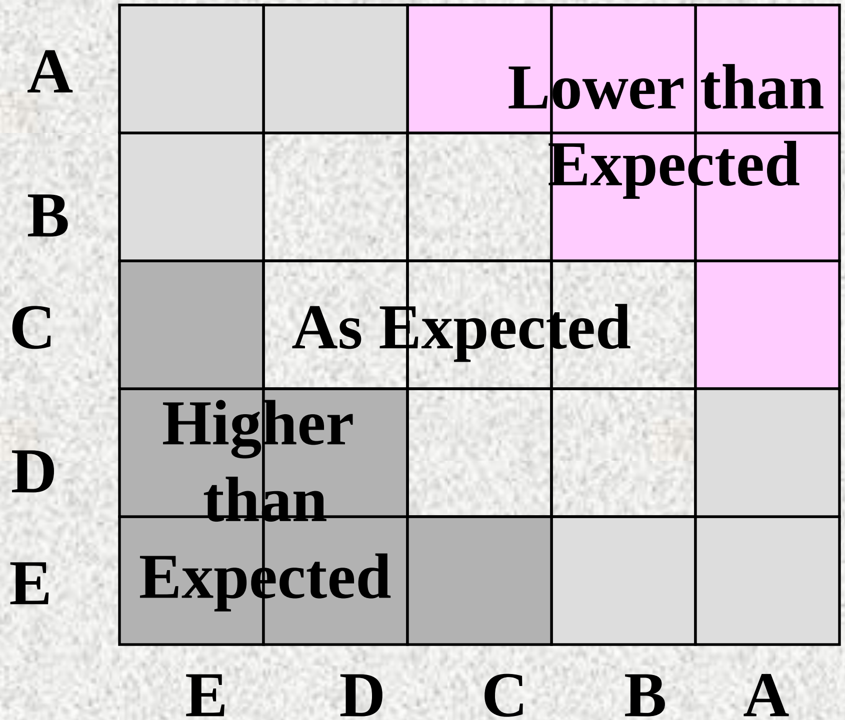
X3 = proportion of water delivered to measured non-households

(Chaplin, United Utilities)

OFWat Trade-offs: Performance Benchmarking (X Factors)

Change in Price Cap: CPI – X + K +Q

OPEX Analysis



Capital Expenditure Analysis

A: excellent
E: very weak

AA: clearly outstanding

CPI: Consumer Price Index
X: Productivity Adj.
K: Capacity Adjustment
Q: Environmental Adj.

Example: NVE Benchmarking

Cross-Sectional Comparisons(200utilities)

Changes over Time

- Physical Quantities (kWh delivered, wires, share low voltage, # customers)
- Cost Structure (endogenous/exogenous, cost components’ shares)
- Cost per physical unit

League Tables--public information/caps

Jan Moen: **Norwegian Hydro Resources and Energy Admin.**

Step 4. Conduct Consistency/Sensitivity Tests

Accuracy and robustness of inefficiency estimates are important due to financial or social impacts.

- Cost Function vs. Production Function
- Functional form (linear, nonlinear)
- Outputs and inputs (e.g., network length vs. fixed assets)
- Alternative methodologies (e.g., DEA vs. SFA).

Need to check whether estimated inefficiency scores or rankings are sensitive to the benchmarking method,

Step 5. Develop Policy Implications

Analyze scores and rankings

Explore in greater detail the potential determinants of inefficiencies across firms and over time.

Firms should not be ranked as poor performers if they operate under conditions that differ from those of the other firms.

Identify the impact of factors like region, population density, regulatory environment, ownership structure, and network vintage.

Seek public comments.

Design incentives to improve performance.

Creating Appropriate Yardsticks

Regulators want to induce outcomes comparable to those achieved under competition.

Pass-Fail Standards?

Reward outstanding performance

Penalize weak performance

Benchmarking provides Yardsticks

Improving Health: “Do No Harm”

- Benchmarking specialists produce and critique studies that utilize various methodologies.
- Rankings can be manipulated by choice of variables, model specification, sample size, time frame, and treatment of outliers.
- Results can be misinterpreted and misused.
- The stakes are high, since affected parties have an interest in the relative and absolute performance comparisons prepared by analysts.

Benchmarking is Part of the Managerial and Policy Tool-kits

The application of the techniques summarized here can improve service quality, expand networks, and optimize operations.

Any benchmarking study will have limitations, but sound studies can be used to place the burden of proof on other parties who might argue that the analysis is incomplete or incorrect.

Over time, data availability will improve and studies will be strengthened as professionals gain experience with these quantitative techniques.

Summing Up

Rankings can serve as catalysts for better stewardship of water and other resources.

If regulators cannot identify historical trends, determine today’s baseline performance, and quantify relative performance across utilities, then as an Indian regulator said, they may as well be writing “pretty poetry”.

For detail, see

Sanford Berg, *Water Utility Benchmarking: Measurement, Methodologies, and Performance Incentives*, International Water Association 2010.