

Advances in Benchmarking to Improve Water Utility Operations: A Review of Six IWA Books

by Sanford V. Berg, University of Florida

May 21, 2012

Abstract

Six recent publications by the International Water Association underscore the growing importance of quantitative studies of water utilities for identifying (1) performance patterns, (2) trends in key indicators, and (3) best practice. Since each author takes a different cut through the problem of performance measurement and incentives for improvement, the books provide different perspectives on the decision-relevance of empirical analysis and cases for managers and policy-makers. Key lessons from the volumes include the use of benchmarking as a tool for establishing better internal incentives, the importance of data collection and auditing, and the role of transparency in helping stakeholders become more informed about utility operations. Analysts need to help stakeholders understand sector complexities, where hydrology, topology, historical developments outside current management's control, population income levels, and the amount of skilled labor differ widely across and even within nations. The six books taken together underscore the value of quantifying relative performance and improving our awareness of production processes and the institutional arrangements that support strong water sector performance.

As has been noted before in the context of evaluating water utilities, if we have no data on performance trends, lack authoritative comparisons among current operations, and have no idea of best practice, decision-makers cannot design performance incentives or establish reasonable targets for key indicators. Stakeholders need to know about the *past* (where operators have been), the *present* (relative performance today), and best practice (so successful techniques can be emulated in the *future*). A recent survey on the quantitative studies underscored the dramatic growth of published research and the dispersion of publications across several fields of study (Berg and Marques, 2011). Six recent International Water Association books consolidate much of the previous research, illustrating ways quantitative studies can be useful to decision-makers. Since they each take a different cut through the problem of performance measurement and incentives for improvement, each provides a different perspective on the decision-relevance of empirical analysis and cases for managers and policy-makers. The six volumes are briefly summarized below:

Institutional Governance and Regulation of Water Services: The Essential Elements (2007) by Michael Rouse identifies the types of institutional structures that are beginning to show success in developed and developing countries. He lays out the policy issues and establishes principles for promoting high performance – drawing from the experiences of both public and private operators. Rouse pulls together the main themes of his book using the case of England and Wales as evidence of how consolidation, corporatization, benchmarking (yardstick competition), and regulatory incentives for cost containment can yield benefits to consumers. His insights set the stage for a series of volumes that support the principles he emphasizes.

Making Public Enterprises Work—From Despair to Promise: A Turnaround Account (2009) by William Muhairwe focuses on Uganda’s National Water and Sewerage Corporation (NWSC). This book examines the people and processes that transformed a moribund state-owned enterprise into one of Africa’s most successful publicly owned organizations. Muhairwe puts flesh on key managerial principles: having a customer focus, operating strategically in a political environment, divisional (decentralized) incentives, performance monitoring, and stakeholder relationships. He identifies four basic ingredients for organizational transformation: thoughtful leadership, careful measurement, open communication channels, and well-designed implementation strategies.

Water Utility Benchmarking: Measurement, Methodologies and Performance Incentives (2010), by the author of this article, provides an overview of the strengths and limitations of different methodologies for making performance comparisons over time and across water utilities (metric benchmarking). In addition, the book identifies ways to determine the robustness of performance rankings. Partial and total metrics (based on cost and production functions) do not identify specific stages of the process that fall short of best practice, nor do the results of metric benchmarking studies show managers how to implement improvements. Rather, the purpose of this book is to demonstrate how analysts can measure water utility operations (that transform inputs into outputs) to perform company comparisons. Those comparisons can then be used to document the need for infrastructure reform and as the basis for regulatory and managerial incentives.

Benchmarking Water Services: Guiding Water Utilities to Excellence (2010) by Cabrera et. al. avoids the terms *metric* and *process*, supporting the IWA Specialist Group on Benchmarking recommendation that “performance assessment” and “performance improvement” be viewed as consecutive components of benchmarking. The volume draws heavily from the European experience, surveying how engineering techniques have been applied to benchmarking. In this “manual of best practice”, the authors emphasize how data accuracy should be quantified and how different types of data analysis can reveal patterns. The authors offer practical steps for translating assessments into improvements. The volume underscores the importance of project planning, as well as specifying the goals of the project when designing data collection and analysis programs.

Regulation of Water and Wastewater Services: An International Comparison (2010) by Rui Cunha Marques provides an overview of the tasks undertaken by agencies providing regulatory oversight. Regulating service quality, promoting transparency, benchmarking performance, and other activities are described in terms of best-practice. After considering rate of return regulation and regulation based on performance incentives, the author provides case studies of eighteen agencies from developed and developing countries. These authoritative reviews of each country include the historical background, legal and institutional framework, market structure, regulatory rules, and an evaluation of sector performance. The concluding sections examine the public service obligations of utilities and the role of regulation in improving performance.

Utility Benchmarking and Regulation in Developing Countries: Practical Application of Performance Monitoring and Incentives (2011) by Silver Mugisha presents various practices, principles, and examples of performance monitoring and incentive design for utilities in developing countries. Thus, it complements the other volumes by outlining the fundamentals of monitoring utilities, formulating corrective actions, and gauging success. The book draws upon the author’s experience at Uganda’s NWSC (the focus of Muhairwe’s volume), and presents practical techniques for monitoring performance. In addition, Mugisha presents detailed incentive plans from Kenya, Nigeria, Tanzania, Uganda, and

Zambia. These incentive formulas utilize partial metrics (performance standards) described in the other volumes.

The six books have water utility data collection, analysis, and performance improvement in common. However, each book has a different emphasis or draws upon different cases. Rouse describes how benchmarking was utilized in the UK regulatory system (emphasizing the importance of governance. Muhairwe covers the case of NWSC's reforms in Uganda (describing the role of performance targets and incentives). The author's volume provides an introduction to metric benchmarking (input-output analysis using production and cost functions) and outlines the strengths and limitations of different methodologies to non-statisticians. Cabrera et. al. examine performance assessment as it is linked to performance improvement (especially as related to processes in various stages of production). Marques provides an encyclopedic review of water utility regulation in practice. Finally, Mugisha focuses on the importance of incentive schemes that promote strong performance.

The books will appeal to a wide range of readers. Economists and Operations Researchers might be most comfortable with elements in the volumes by Rouse and this author, with the emphasis on governance, metrics, regulation and incentives. Managers will find Muhairwe's detailed case study full of insight and inspiration. They will also be comfortable with the approach in the Cabrera et. al. book that identifies best practice for benchmarking. Marques' compendium of regulatory oversight around the world underscores the role of government regulation (and associated rules and incentives) on improving utility performance—providing policy-makers and regulatory professionals with a comprehensive resource. Finally, Mugisha's translation of principles into practice will serve as a Bible for those who are trying to use performance monitoring to design specific incentive schemes for water utilities in developing countries.

Is it possible to identify a policy consensus based on these six disparate books? Taken together, they support a position that sound engineering is necessary but not sufficient for high performance. What follows is this author's distillation of key lessons from the authors—all of whom are experts in their specialized fields. Of course, this is one individual's view; and as author of one of the volumes, my judgment regarding the contributions of the books should be taken with a grain of salt. Nevertheless, the following points receive solid support in these volumes:

Data analysis involves both subjective and objective elements: Without data, decision-making cannot be based on reality. The subjective aspect of analysis arises from placing weights on the different dimensions of performance (for example, expanding network coverage versus improving service quality). Is non-revenue water a bigger problem than providing service for only eight hours per day? The answer depends on social values, as translated into laws, which determine the weights to be given to different performance goals. The objective component is based on utilizing sound principles for data collection and auditing (Cabrera). It also relates to the application of appropriate methodologies for evaluating performance over time and across decision-making units (Berg).

Benchmarking represents a valuable management tool that enables decision-makers to identify and reward top performing units: "Executives manage what they measure." The absence of data is evidence of weak managerial processes. Of course, not all that is important is easy to measure, but it is better to have an approximation of an important indicator than a precise calculation of an irrelevant indicator (Mugisha). The development of strong internal incentives can be particularly problematic for state-owned enterprises, but examples of success stories support the view that inducements can be devised and implemented (Muhairwe).

Information helps different stakeholder groups understand the performance of different units: The affected communities deserve to be informed of trends and relative performance, so pressure can be placed on those responsible for improving sector performance (Rouse). Good governance requires transparency and citizen access to annual reports that provide financial and operational information. The data ought to be accessible to citizens and to analysts, so that trends can be understood.

Multi-period information on operations and financial conditions is essential for decision-making: Retaining historical data provides analysts with the opportunity to identify trends and conduct more robust statistical analyses. One cannot understate the importance of being clear about the objectives of the project, the techniques being used, and the level of detail required for the dataset (Cabrera). In addition, “partial indicators” (such as labor productivity) should be used in conjunction with other data if the comparisons are to be robust and credible.

One objective of a benchmarking study is to measure productivity and efficiency so that the analyst can make comparisons: *Productivity* considers the link between inputs and an organization’s outputs. *Efficiency* is related to productivity, but it involves establishing a standard and determining how close the firm comes to meeting that standard: how far is the utility from “efficient practice”? How near is the utility to the frontier? Quantitative studies can be used to place the burden on management to explain why outcomes deviate so dramatically from what the “model” suggests is feasible. At the same time, those conducting statistical studies and data envelopment analyses need expertise. At the same time, they should exhibit a high degree of humility when presenting results (Berg).

Benchmarking can trigger political changes internal to organizations: As performance is highlighted on a regular basis, those responsible for implementing successful cost reduction programs are likely to gain influence within the utility (Muhairwe). Greater transparency and public awareness of relative performance put pressure on weak utilities to restructure their management teams or to develop better incentives for meeting well-defined targets. Of course, this puts an additional burden on those conducting benchmarking studies that are utilized for comparative performance assessment (yardstick competition in the context of regulation).

The ability of researchers to conduct and evaluate studies is affected by database limitations and methodological shortcomings: Clearly, information contains errors (whether due to errors in reading meters or transferring data from sources to accounting reports). Managers need to continually seek ways to improve their information systems (Cabrera). Similarly, the various benchmarking techniques each involve limitations—whether the methodology involves partial indicators, statistical estimates, or data envelopment analysis. Within utilities, managers should “avoid information empires.” When data are widely available, infrastructure professionals are in a better position to make better operating and planning decisions.

Demand forecasting is essential for business planning (for investments and operations): Customer usage data and population growth can be used for forecasting likely future demand. Business plans serve as reality checks for decision-makers: are the cash flows reasonable and will the operational targets be met under current financial constraints? Will quality of service be improved under the business plan? Too often in developing countries, political rhetoric establishes unrealistic expectations within the citizenry. Countering this tendency, responsible public officials need to identify specific targets and the resources required to reach those objectives (Rouse).

Benchmarking puts pressure on management and regulators: When citizens see what some utilities (or divisions of a utility) have been able to accomplish, they are in a position to reward strong performance and raise questions regarding weaker performance. Without information, citizens basically operate in a vacuum. Similarly, various groups have performed benchmarking studies of regulatory agencies, focusing on both adherence to accepted (and mandated) processes and actual performance in the sector being regulated (Marques).

Benchmarking is a key instrument for promoting transparency and access to information: Annual reports and other communications from utilities provide some data that can be used for making comparisons. However, examination of web sites and actual annual reports indicates that many utilities fall short of transparency: some web sites are quite uninformative and many annual reports lack audited financial data and are not issued in a timely manner. Thus, it is important to have more comprehensive studies that focus on performance over time and across comparable decision-making units (Berg).

Benchmarking reduces information asymmetry for decision-making: Within a utility, executives at the highest levels tend to receive summaries of information, yet the detail can reveal patterns that are very important. Without more careful data analysis, the executives are in no position to develop and implement new strategies for improving performance (Muhairwe). Similarly, regulators have far less information on actual and potential performance than managers. Getting the information out into the public domain improves decision-making: “People are entitled to their own opinions, but not to their own facts.”

Benchmarking sets baselines and facilitates evaluation: Managers need a clear idea regarding the actual status of operations and the financial sustainability of their utilities. Benchmarking documents current conditions and performance, so improvements can be quantified relative to current indicators. In addition, there are a variety of techniques utilized around the world summarized in several of the books: the American Water Works Association QualServe Program, the UK’s OFWAT models, regional studies conducted in South East Asia and in Latin America, and IB-Net’s data collection initiative. Some of these programs examine broad patterns, while others focus on improving specific functions or stages in the production process. Ultimately, it is important that managers devote resources to specific improvement actions and to reviewing outcomes (Cabrera).

Benchmarking can be used to establish targets: The best-performing decision-units provide evidence of best practice: the current frontier. How quickly other decision-units should be able to move to that frontier is a more complicated question requiring input from all stakeholders, including utility management. However, with benchmarking, the burden of proof is placed on decision-makers to explain their organizations’ relative performance and their ability to move to the frontier (Mugisha).

Benchmarking promotes continuous improvements: Key Performance Indicators (KPIs), like output per worker, can be used to track improvements in utility performance. Another type of benchmarking, called *process benchmarking*, focuses on individual production processes in the vertical production chain. Managers from different organizations can share “best practices” and improve utility performance (Cabrera). One advantage of this approach is the ability to identify specific stages of the production process that warrant attention (including pumping, intake, transport, clarification and filtration of groundwater, purification and treatment of raw surface water, maintenance, meter reading, collections, planning, etc.).

Metric benchmarking includes a variety of quantitative tools: Empirical studies can play a role in identifying relative performance. The various methodologies have different degrees of sophistication (each with their associated strengths and limitations). Key Performance Indicators (KPIs) tend to be partial indicators, capturing one feature of performance. Statistical methods include Ordinary Least Squares (OLS), Corrected Ordinary Least Squares (COLS), and Stochastic Frontier Analysis (SFA). These methods are labeled *parametric*, since they provide statistical measures of goodness of fit and parameters (or coefficients) for the cost or production functions. Data Envelopment Analysis (DEA) is a *nonparametric* technique in which linear programming is applied to a selected set of variables to calculate an efficiency score for each utility (or decision-unit). DEA has the advantage of not requiring a specific functional form. Metric benchmarking can identify outliers—providing managers and regulators guidance on which utilities have the strongest and weakest performance (Berg)

Reports should avoid the use of excessive jargon: Ultimately, the results of studies need to be communicated to non-specialists. Technical terminology has its place, but only for technical communities. There are different audiences for benchmarking reports. The analyst should know the audience (a legislative committee or the general public) and determine the appropriate format for the report. Such reports can be misinterpreted or misused, so great care must be taken to explain how robust the results are and how the ranking will be utilized. Preliminary reports can seek input from stakeholders. Final reports should be widely available for review (Berg).

The focus should be on what is important, not what is easily measured: Studies can hide as much as they reveal. Analysts should address the real issues and not bury the implications of a study in footnotes. Readers will want to understand why the study was conducted and how it will be used. An authoritative and robust study should point out the strengths and limitations of the methodologies utilized *and* the implications for decision-makers. If the study is preliminary in nature, that should be made clear early in the report. Finally, “The perfect is the enemy of the good.” No study is perfect. The task is to get started (Mugisha and Marques).

Benchmarking can be disruptive: The press is likely to sensationalize the comparisons and vilify managers and regulators. Comparisons are bound to create problems for those utilities identified as “weak” or for those divisions (within a utility) identified as performing poorly. People will point fingers and managers will scramble. When dramatic change is called for, such disruption is quite appropriate. However, the analyst needs to be completely candid about his or her confidence in the results (or rankings) determined by the analysis. Managerial and political careers can be affected by performance benchmarking studies (Rouse and Cabrera).

Benchmarking studies force decision-makers to face reality: Facts focus attention on particular dimensions of performance. So long as there are no facts, different groups can make claims, but those claims will not be backed up by evidence. Benchmarking changes the “game” since it provides information to stakeholders. That information should capture what is and what is not within management’s control: both the context within which the industry operates and the actual performance of service providers. Rhetoric about water or energy as a “human right” that should be available to all can help citizens prioritize sectors receiving public funds. However, that rhetoric is empty if it does not specify a plan for achieving specified goals (Marques and Muhairwe).

Analysts should draw upon the skills and support of personal networks of experts: No analyst can afford to work alone. The preparation and dissemination of sound benchmarking studies require a wide range of skills, some technical and some strategic. Continuing education is essential if analysts are to

remain abreast of new developments. Colleagues in organizations around the world are able to provide advice and feedback on technical studies. Analysts can work with local university researchers and with professionals who share their interest in strengthening performance in infrastructure sectors (Cabrera).

None of the points presented above should be taken as an “absolute”. We are dealing with a very complicated sector where hydrology, topology, historical developments outside current management’s control, population income levels, and the amount of skilled labor differ widely across and even within nations. IB-Net underscores the wide diversity of operating conditions. However, the six books taken together underscore the value of quantifying relative performance. The authors have contributed to our understanding of production processes and the institutional arrangements that support strong water sector performance. Yet in a sense, we are still in the early stages of a renaissance in the way water systems are governed, incentivized, and operated.

References

Berg, S. and R. Marques (2011). “Quantitative studies of water and sanitation utilities: A benchmarking literature survey,” *Water Policy*, 13, 591-606.

Berg, Sanford (2010). *Water Utility Benchmarking: Measurement, Methodologies and Performance Incentives*, International Water Association, xii-170.

Cabrera, Enrique Jr., Peter Dane, Scott Haskins, and Heimo Theuretzbacher-Fritz (2010). *Benchmarking Water Services: Guiding Water Utilities to Excellence*, International Water Association, xix-164.

Carvalho, Pedro, Rui Cunha Marques, Sanford Berg (2012). A meta-regression analysis of benchmarking studies on water utilities market structure, *Utilities Policy*, 21, 40-49.

IB-Net: International Benchmarking Network for Water and Sanitation Utilities, <http://www.ib-net.org/>.

Marques, Rui Cunha (2010). *Regulation of Water and Wastewater Services: An International Comparison*, International Water Association, xix-292.

Mugisha, Silver (2011). *Utility Benchmarking and Regulation in Developing Countries: Practical Application of Performance Monitoring and Incentives*, International Water Association, xi-195.

Muhairwe, William (2009). *Making Public Enterprises Work—From Despair to Promise: A Turnaround Account* International Water Association, xiv-415.