Water Utility Benchmarking for Managerial and Policy Decisions: Lessons from Developing Countries

By S. V. Berg[1] and M. L. Corton[2]

December 2007

Prepared for the International Water Association Pi08 Conference on Performance Assessment of Urban Infrastructure Services (March 2008, Valencia, Spain)

Abstract

Those responsible for utility operations can only manage what they measure, so having information on productivity trends and relative performance enables utility managers to direct attention to shortfalls. Similarly, policy-makers require quantitative analyses in order to identify utilities with strong and weak performance. Examples from Central America are used to illustrate key points.

Keywords: Benchmarking, DEA, SFA, Performance

1. Introduction

A recent IADB study reports that investments of $40 billion for water assets are needed to meet the United Nation's Millennium Development Goals; wastewater treatment would significantly raise that funding requirement ("Obstacles and Constraints, 2003). A survey of 400 stakeholders included in the study identified inappropriate pricing policy and lack of clarity in regulatory processes as the two major constraints for increasing investment in water and sanitation systems (WSS) in Latin America and the Caribbean. Private sector funding could play a role in expanding or improving urban water systems through either equity investments or the purchase of municipal bonds. However, external financial flows are required to increase significantly absent major improvements in system performance and to develop incentives for better WSS performance – in the sense of increased quality of service provision and cost efficiency of the utility system's operations.

Expecting infrastructure investment to grow in Central America, the Inter-American Development Bank (IADB) funded a PURC study on *Benchmarking Water Utilities* in the region. The aim of IADB was to gauge the impact of loans on network expansion (coverage) and on service quality. In addition, water professionals at international organizations must be able to understand what utilities (and nations) are doing as "best practice" so that incentives can be developed to enhance performance (Baietti, et. al., 2006). The project involved conducting metric benchmarking analysis using data gathered from national regulators and utilities. This work complements and extends the Association of Water Regulators of the Americas' (ADERASA) recent empirical study on benchmarking Latin American water service providers.

[1] *Distinguished Service Professor, Economics; sberg@ufl.edu, Public Utility Research Center, University of Florida, Gainesville, FL 32611,+1 352-392-0132 Fax +1 352-392-7796.*

[2] *PURC Research Associate, mcorton@ufl.edu, Public Utility Research Center, University of Florida, Gainesville, FL 32611, +1 352-392-6148 Fax +1 352-392-7796.*

2. Lessons

To some extent, academic research tends to be technical—drawing upon advanced statistical and data analysis techniques. Studies useful for decision-making tend to be more intuitive. The latter attempt to translate principles into practice, but excessive emphasis on simplicity can be a problem. Utility managers and regulators in Central America reviewed the PURC research at the recent PURC/IADB Benchmarking Workshop (San Jose, Costa Rica, October 15-16). Participants from six nations identified factors having an impact on data quality, data collection, benchmarking methodologies, and possible policy and regulatory implications of performance rankings. In addition, lessons from academic research underscored the need for sensitivity tests before utilizing "scores" or rankings for setting prices or rewarding managers.

2.1 Lessons from Managers: Keep it Simple

The Workshop provided two opportunities for practitioners to respond to technical studies and to issues raised by regulatory policy-makers. The reflections have been distilled into ten key points, but capturing all the nuances presented by various individuals is beyond the scope of this study.

2.1.1 Data Manager Commitment

The managers responsible for data collection, verification, storage, and processing must be convinced that this activity is vital to the success of the enterprise. Thus, there must be an understanding between the operator and the regulator. The operator should be convinced that the task goes far beyond fulfilling a requirement of a financial statement where numbers are placed in particular columns, but also about how the data is useful for management. Besides serving as a report to the regulator or to any external institution, data must be viewed as important and useful for the company – for strategic, operational, administrative and commercial purposes. Benchmarking is a tool that results in resource savings for the company concerned with performance: the information system pinpoints the areas that should be targeted for future initiatives and allows managers to evaluate the impacts of past interventions.

2.1.2 Data Manager Continuity

There needs to be a high level person responsible for data within the company; however, it is the position (rather than the person) that must have continuity over time within the company. This formalized role is needed to address internal turnover problems which limit data collection, causing gaps in time series and in cross-section observations. Obtaining data series for a number of years also requires the support of funding agencies, including national development banks. The lack of data in systems like IBNET (the World Bank's data library for water utilities) is a sad commentary on past incentives. It suggests that national development agencies and international banks have only been focusing on benchmarking in recent years, perhaps because cutting ribbons for new projects is much more glamorous (and less threatening) than evaluating past decisions.

2.1.3 Centralized vs. Decentralized Systems

Centralization and decentralization are both viable strategies, but the strengths and limitations of each must be recognized as nations and organizations adapt to changing conditions. To some extent, if a system is "working well" there is no need to dramatically change institutional arrangements. Nations and organizations have different historical contexts; the development of new sector policies is highly path-dependent. Thus, within an organization, duplication of data storage files inside the firm and data reports (in specialized formats) to external institutions raises administrative costs. This would suggest centralization—especially given the opportunities associated with new information technologies. On the other hand, centralization without access opens up the possibility for little "Information Empires," where individuals exercise power by withholding data from those who should have access to information. Balancing such considerations leads to different types of policy initiatives. For example, Honduras has gone through a decentralization process (Pearce-Oroz, 2006) while other nations have consolidated their utility suppliers. There is no single recipe for promoting high performance: rather, leaders build upon the experiences and institutions that provide the baseline when making comparisons.

2.1.4 Data Disaggregation

Data disaggregation improves decision-making. Clear customer classification (residential, industrial, and commercial) allows for more accurate information regarding operation and performance of the company. In addition, maintaining data series on particular regions or divisions of a company allows top managers to develop strategies for rewarding strong performance. Disaggregated data allow managers to target areas of sub-standard performance and facilitates quantitative studies of cost and productivity. For example, Mugisha et. al. (2007) show how Uganda's state-owned NWSC utilizes disaggregated benchmarking information to set targets, compare managerial performance, and determine management bonuses (up to 50% of base pay).

2.1.5 Data Definitions and Benchmarking Objectives

Clear variable definitions allow outsiders to interpret information; consistency and clarity are fundamental to the management process. The International Water Association has played a leadership role in this process.

2.1.6 Operational Data

Better operational data collection procedures are needed: timely reports that identify patterns mean that network repairs can be addressed in a comprehensive and cost-effective manner.

2.1.7 Other Data Sources

Workshop participants urged that data collection not stop at items under the firm's direct control. For example, rural systems must still be monitored—by a government ministry or a regional water utility (not currently serving those potential customers). One representative was adamant that social indicators be included in studies, so that service quality and continuity could be linked to socio-economic characteristics of different regions. Furthermore, factors external to the company affect production costs and opportunities for productivity growth. Distance from water source, geographic characteristics (topology and population density), and a culture of payment (or non-payment) all affect costs. External factors also may have an impact on the collection and storage of data. For example, the existence of records and maps of the city; the frequency of the country census; municipal or city restrictions regarding the network design; number of connections per km, and type of users of the network. Thus, empirical researchers need to take into account factors beyond managers' control just as managers need to draw upon the census and other external data-sets in their own system-planning.

2.1.8 Information Technology and Management

Information technology is necessary, but not sufficient, for sound management: information systems should link financial-commercial-operational data. Leaders can only manage what they measure. Aid projects that computerize companies without altering underlying internal incentives cannot have significant impacts on performance. The entire supply chain would be better off if utilities moved more quickly towards best practice: those manufacturing pipes, pumps and other hardware would see a dramatic increase in business in particular countries *if* utilities demonstrated good stewardship.

2.1.9 Transparency and Public Policy

The company information needs to be public to promote managerial accountability and citizen confidence in infrastructure services. Some state-owned enterprises become the home of political patronage. When there are poor internal incentives, the performance of an entity is likely to be sub-standard. If customers (and un-served citizens) do not have data on comparable utilities, the citizens are in no position to put pressure on managers to improve performance. There is evidence that even rough comparisons can put pressure on political leaders to fulfil promises to provide funds for network expansion and on managers to deliver services at least-cost (Rossi and Ruzzier, 2000). However, regulatory commissions and/or water ministries must be committed to institutional reform if they are to risk offending powerful groups (including unions and high income customers now being subsidized).

2.1.10 Methodologies and Capacity-Building

Overall, there is a need within the firm for more information regarding benchmarking methodologies and their application. Larger water utilities have engineers who are familiar with process benchmarking. There is also a need for capacity building in the area of metric benchmarking—starting with trends in Core Indicators, and moving to basic statistical reports and DEA studies. Core indicators include labor productivity, water losses, and trends in average cost. Statistical studies can identify cost drivers and determine how far utilities are from the production frontier.

2.2 Lessons from Academics: Recognize Complexity

The last point from managers provides a natural link to those conducting more technical analyses. Managers are quite aware of the different conditions facing different water utilities and how difficult it is to make valid comparisons across firms and over time. Thus, they may be sceptical of using rankings to reward utilities, set prices, or to reward divisional managers. Nevertheless, information on broad patterns allows political leaders and citizens to "grade" the performance of water utilities.

2.2.1 Sources of Productivity Change

Performance indicators are "partial" in nature: they consider one dimension of performance at a time. By contrast, total factor productivity (TFP) analysis focuses on firms productivity change over time and takes several inputs and outputs into account. Productivity may differ over a period of time for the following reasons:

a. Technological change (frontier shift): a technological change has occurred within the sector producing a variation on firm's productivity.
b. Technical efficiency change (catch-up): A firm has varied its production efficiency over time.
c. Scale efficiency change: A firm change in production efficiency is product of a change in its scale of production.
d. Input mix allocative efficiency: A firm has varied its efficiency by producing outputs using a different mix of inputs, given the input prices the firm faces;
e. Output mix allocative efficiency: A firm has varied its efficiency by producing a different level of outputs with the same mix of inputs; given the input prices the firm faces.

By understanding the sources of productivity change, managers can focus improvement in areas that seem weak. At the same time, by understanding the sources of productivity change, policy makers, investors and other stakeholders can point to the most productive firms as examples of strong performance—promoting the diffusion of best practice to all firms.

To decompose productivity into the above components requires knowledge of the firm's production technology – the particular way a firm mixes inputs to produce its set of outputs which generally translates into a functional mathematical form for the production process. Quantitative methodologies can be used to isolate the roles of these various components which are generally accomplished by employing econometric procedures to estimate a functional form for the production technology. Statistical tests are performed to check how properly represented is the technology by the selected functional form (Garcia and Thomas, 2001).

2.2.2 Strengths and Limitations of Methodologies

A single index of water utility performance has the same problems of any indicator: it will be neither comprehensive nor fully diagnostic. The ratio of water delivered to number of workers does not capture water losses or the sustainability of operations when certain types of expenditures (like maintenance) are deferred. Nevertheless, empirical studies of water utilities within and across countries have yielded insights into the determinants of relative performance by water utilities (Cubbin, 2005).

The development of sound managerial and public policy incentives involves data collection, analysis, and interpretation. Several methodologies should be used to determine the sensitivity of the results (and performance rankings) to model specification and outliers. Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA) are two approaches to production and cost functions. The limitations to such studies must be balanced against the information asymmetries that arise in the absence of such studies.

The methodologies used by different groups differ in ways that are unsurprising. Practitioners seek decision-relevant studies that will help them document best practice and justify actions (and investments). Policy-makers tend to step back from the actual operations of firms, and seek to document the impacts of policies on sector performance and to identify troublesome patterns. Regulators, in particular, attempt to design incentive plans that utilize relative performance information. Academic researchers, on the other hand, tend to look for sophisticated ways to address more technical issues: the extent of economies of scale, the rate (and direction) of technological change, and sources of inefficiency. Each of these areas of concern justifies some degree of specialization and division of labor. As we shall see, each tends to use different types of benchmarking studies. However, the walls separating practitioners, policy analysts, and academics are higher and thicker than they need to be.

2.2.3 Accessibility of Technical Studies to Practitioners

The reports developed by regulatory commissions and those published by academic researchers have different purposes. Bridging the two types of studies requires one foot in the world of politics and utility operations and one foot in the world of advanced data analysis: few of us have such skill sets. That suggests that teams of analysts from both worlds might have the greatest chance to obtain needed data, ask the right questions, and interpret the results in ways that are intuitively understandable to those without advanced skills. Such studies would include sensitivity tests, so that the results are robust—providing external observers with some confidence in the conclusions. The PURC/IADB Project for Central America (Corton, 2007) represents a starting point for a type of collaboration that shows some promise—as it engaged practitioners in the research process.

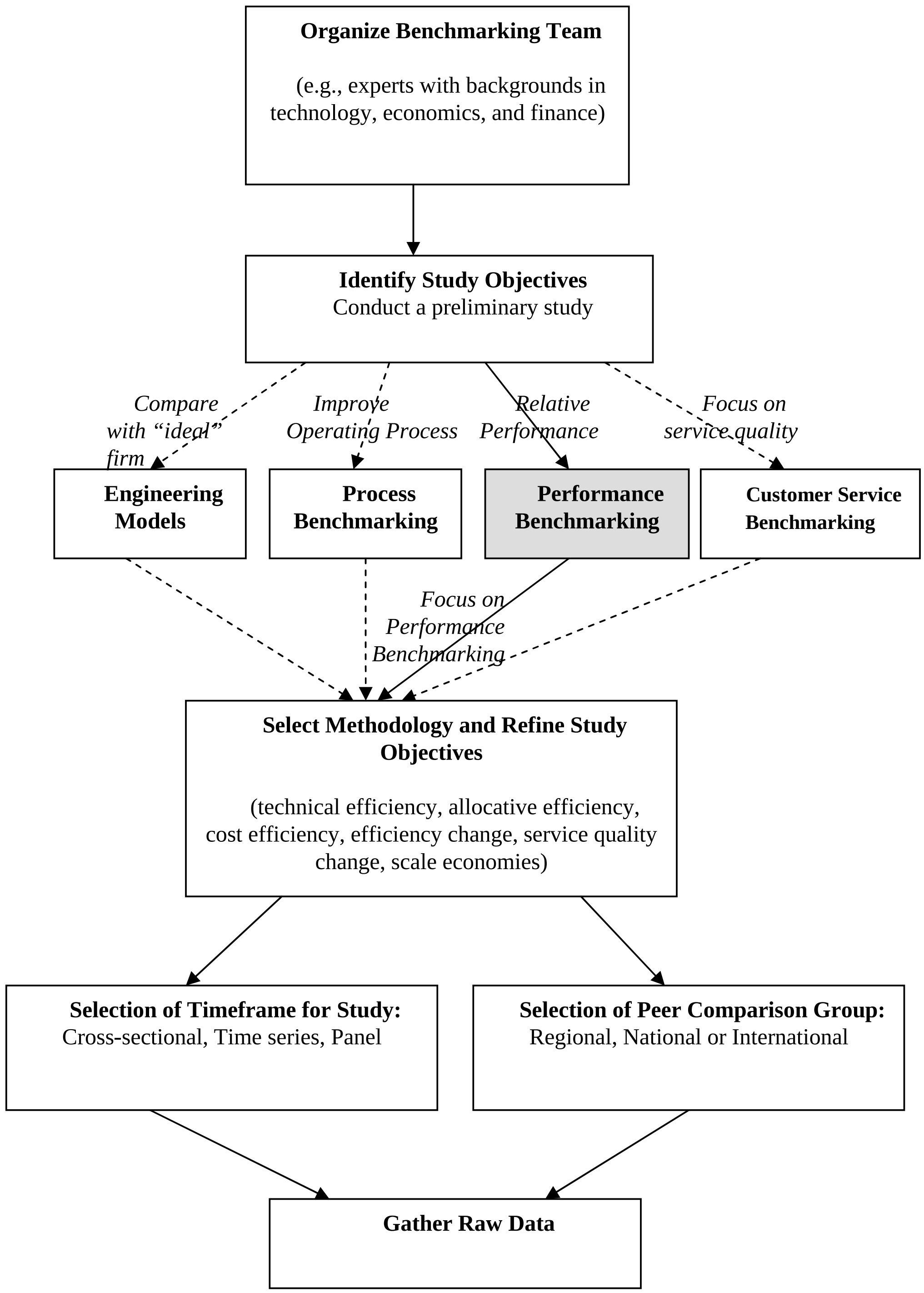
3. Political Economy of Performance Indicators

The key lessons regarding collecting and analyzing operating and financial information underscore the practical orientation of decision-makers. Practitioners, policy-makers, and academics all have studied water sector performance. The three groups have tended to operate in relative isolation. Given the important resource allocation issues associated with infrastructure industries in developed and developing nations, greater communication among the three types of researchers is called for. However, such cross-fertilization requires that each group understands the priorities placed on different aspects of inter-firm comparisons. The benchmarking process can be divided into five steps: (1) identify objectives, select methodology and gather data; (2) screen and analyze data; (3) utilize specific analytic techniques; (4) conduct consistency/sensitivity tests; and (5) develop policy implications.

3.1 Steps in Benchmarking

Here, we focus on the first step: objectives and methodologies. Figure 1 depicts the elements and shows four broad methodologies. In addition, several other types of performance indicators must be incorporated into comprehensive benchmarking studies.

Figure 1
Step 1: IDENTIFY OBJECTIVES, METHODOLOGY, AND DATA



The first step in the benchmarking process requires the policy-maker or analyst to identify issues to be addressed, the time period to be analyzed, and the types of comparisons to be made. These choices will reflect current analytic capabilities, an initial understanding of data availability, and preliminary methodological choices. The objectives of benchmarking studies tend to depend on emerging policy issues: whether sector consolidation (or disaggregation) might promote greater efficiency, whether privatization is associated with performance improvements (Estache and Rossi, 2002), and which managerial strategies are most effective in containing cost or improving service quality. As Figure 1 indicates, there are a number of alternative approaches to benchmarking, including engineering models, process benchmarking, performance benchmarking, and customer service benchmarking. Regulators utilize all four approaches to different degrees: Chilean regulators use the hypothetical firm (engineering) approach to establish performance objectives, managers tend to focus on specific processes (to identify best practices), regulators everywhere utilize performance indicators to evaluate utilities, and managers conduct customer surveys to determine areas for improvement. The focus here will be on the third category of benchmarking indicators based on performance studies that utilize cost or production functions. Such studies have their own strengths and limitations, but these methodologies provide a valuable framework for investigating a number of questions.

3.2 Types of Indicators: Decision-Relevance

3.2.1 Core Overall Performance Indictors

Core Overall Performance Indicators include a number of *Specific Core Indices*, such as volume billed per worker, quality of service (continuity, water quality, complaints), unaccounted for water, coverage, and key financial data (operating expenses relative to total revenues, collections). These partial measures are generally available, and provide the simplest way to perform comparisons: trends direct attention to potential problem areas. Policymakers often combine the specific core indices to create an *Overall Performance Indicator (OPI)*, generally using a weighted average of core indices. Thus, an OPI provides a summary index that can be used to communicate relative performance to a wide audience. Regional associations of water regulators have promoted projects that facilitate comparisons of specific core indicators.

3.2.2 Performance Scores: Production or Cost Estimates

Performance Scores based on Production or Cost Estimates are used to identify the best performers and the weakest performers in a group of utilities. The metric approach (favored by academics) allows quantitative measurement of relative performance (cost efficiency, technical/engineering efficiency, scale efficiency, allocative efficiency, and efficiency change). Performance can be compared with other utilities at a point of time and over time, using statistical and/or nonparametric frontier methods. Thus, performance scores and relative rankings identify under-performing and high-performing utilities. Rankings can be based on production relationships and/or cost structures (Berg, 2007).

3.2.3 Engineering/Model Company Approach

The Engineering/Model Company approach has been used to establish baseline performance. This methodology requires the development of an optimized economic and engineering model: based on creating an idealized benchmark specific to each utility—incorporating the topology, demand patterns, and population density of the service territory. The use of an "artificial" firm that has optimized its network design and minimizes its operating costs can provide insight into what is possible if a firm is starting as a Greenfield Project. As with any methodology, this approach also has its limitations. The engineering models that support it can be very complicated, and the structure of the underlying production relationships can be obscured through a set of assumed coefficients used in the optimization process. Chile and Argentina have used this approach for establishing infrastructure performance targets; however, it is not widely used by policymakers, managers, or economists.

3.2.4 Process Benchmarking

Process Benchmarking focuses on individual production processes in the vertical production chain. Managers actively utilize this technique because it can identify specific stages of the production process that warrant attention. For example, to obtain finished drinking water involves the following steps: pumping up, intake, transport, clarification and filtration of groundwater as well as the purification and treatment of raw surface water. Detailed examination of production facilities and their operations would be the starting point for process benchmarking. Similar studies would be performed for distribution processes (network design, pipeline construction and maintenance), sales processes (including meter reading, data processing, billing, collections, and customer relations), and general processes (like planning, staff recruitment and retention, and public relations). Many water associations focus on process benchmarking as a mechanism for identifying potential benchmarking partners, preparing for and undertaking benchmarking visits, and implementing best practices (Larsson, et. al, (2002)). From the standpoint of public policy, there must clear delineation of utility obligations and regulatory responsibilities so that process benchmarking studies do not lead to undue regulatory interference with managerial decision making.

3.2.5 Customer Survey Benchmarking

Customer Survey Benchmarking focuses on the perceptions of customers as a key element for performance evaluation. Unlike the other approaches, this technique can shed light on consumer concerns, reflected in complaints or captured in customer surveys. One widely-used model 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), and empathy (giving individual care and attention). Utility managers conduct such surveys to reveal performance trends over time (Parasuraman, Zeithaml and Berry (1985)). Proactive managers who disaggregate complaints by type of customer, location, and type of complaint can identify problem areas.

3.2.6 Other Indicators: Financial and Resource Sustainability

At least two other indicators shed light on other dimensions of sector performance. Financial sustainability is not captured in other indicators. In fact, a utility could defer needed maintenance and reduce operating costs—possibly yielding lower revenue requirements than other firms. Much more complete models would be needed to test for such problems. Similarly, water resource sustainability is not captured in the other indicators, unless a process indicator includes water sources as a factor.

4. Concluding Observations

The specific coefficients of the Central American Water Utility Benchmarking Study are less important than the fact that another region of the world is engaged in a more deliberate process for making performance comparisons. Analysts apply these quantitative techniques to determine relationships among variables: for example, utilities that produce far less output than other utilities (who are using the same input levels) are deemed to be relatively inefficient. Similarly, a utility might have much higher costs than expected (based on observations of others producing the same output level but having lower costs). A finding of excessively high costs would trigger more in-depth studies to determine the source of such poor performance.

This project establishes a strong case for more comprehensive studies in the future—helping to set the stage for creating strong incentives for improved performance. The water regulatory authority can be instrumental in spearheading many of these activities. However, benchmarking will be much more effective if two points are met. First, regulated companies must cooperate with regulators and subscribe to improvement objectives and processes. Second, formal mechanisms need to be established for consumers and other affected parties to raise concerns and suggest modifications to the process. Through incentive regulation and an appropriate price cap formula, regulators can use findings from benchmarking reports to reward high performance companies (Estache and Kouassi, 2002). They can also pressure laggard companies to promote cost-containment and the improvement of service quality. A properly designed benchmarking system should prevent poorly performing companies from increasing prices as much as the "average" water utility to which they have been compared—so long as prices, do indeed, cover cots. If companies operate more efficiently, customers will benefit from lower prices and should continue to expect and receive high quality service. The resulting system is likely to be sustainable—promoting further network expansion and the adoption of best practice by most water utilities.

REFERENCES

Baietti, A. W. Kingdom, Meike van Ginneken (2006). “Characteristics of well-performing public water utilities,” Water Supply & Sanitation Working Notes No. 9, May.

Berg, S. (2007). “Survey of water utility benchmarking studies,” PURC Working Paper.

Corton, M. L. (2007). “Benchmarking water utilities in Central America,” PURC Working Paper.

Cubbin, John (2005). “Efficiency in the water industry,” *Utilities Policy*, 13 (4): 289-293.

Estache, A. and E. Kouassi (2002). “Sector organization, governance and the inefficiency of African water utilities,” World Bank Policy Research Working Paper No. 2890, Washington, D.C.

Estache, A. and M. Rossi (2002). “How different is the efficiency of public and private water companies in Asia?” *The World Bank Economic Review* 16(1):139-148.

Garcia, S. and A. Thomas (2001). “The structure of municipal water supply costs: application to a panel of French local communities,” *Journal of Productivity Analysis* 16:5-29.

Larsson, Mats, Renato Parena, Ed Smeets and Ingrid Troquet, (2002). “Process Benchmarking in the Water Industry: Towards a Worldwide Approach,” International Water Association.

Mugisha, S., S. Berg, and W. T. Muhairwe (2007). “Using internal incentive contracts to improve water utility performance: the case of Uganda’s NWSC,” *Water Policy*, 271-284.

“Obstacles and constraints for increasing investment in the water and sanitation sector in Latin America and the Caribbean: Survey,” IADB, Dec 2003, 1-13.

Parasuraman, Zeithaml and Berry (1985). “A conceptual model of service quality and its implications for future research,” *Journal of Marketing* 49 (4), Fall, 41-50.

Pearce-Oroz, Glenn (2006). “The viability of decentralized water and sanitation provision in developing countries: the case of Honduras,” *Water Policy*, 8, 31–50.

Rossi, Martin Antonio and Christian Alejandro Ruzzier (2000). “On the regulatory application of efficiency measures,” *Utilities Policy,* 9, 81-92.