BENCHMARKING CENTRAL AMERICAN WATER UTILITIES

Sanford Berg

Public Utility Research Center
University of Florida
sanford.berg@cba.ufl.edu
Phone: (352) 392-0132

Correspondent author: Maria Luisa Corton

Public Utility Research Center
University of Florida
PO Box 117142
Gainesville, FL 32611-7142
Phone: (352) 214-5132
mcorton@ufl.edu

Abstract

This study provides a comprehensive efficiency analysis of water service providers in six countries in the Central American region. Pressures for sector reform have stimulated interest in identifying and understanding the factors that can contribute to network expansion, improved service quality, and cost containment. The aim is to provide policymakers and investment funds institutions with quantitative evidence on the effectiveness of the regional water sectors and utilities under different perspectives. In addition to key sector performance indicators, the analysis considers several benchmarking methodologies to assess technical and cost efficiency.

Keywords: Performance Indicators, DEA, Total Factor Productivity, Benchmarking

JEL Classifications: D24, L23, L51, L95
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.¹ Public and 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 unlikely to increase significantly absent major improvements and incentives in measuring system performance.

Developments over the past decade in quantitative techniques and pressures for sector reform have stimulated interest in identifying and understanding the factors that can contribute to water and sanitation systems network expansion, improved service quality, and cost containment in the sector (Saleth and Dinar 1999). Policymakers in Latin America, Asia, and Africa have begun to collect data that can serve as the basis for performance comparisons that help decision-makers identify weak and strong performers. Utility managers, water associations, regulators, and other groups have begun to undertake statistical analyses of water systems over time, across geographic regions, and across countries.

The purpose of this study is to analyze the relative performance of water utilities in the Central American region to identify best performers and areas of weakness in the sector. The results can help decision-makers’ better direct investment funds into projects that will further develop the water sector in this region. A substantial body of technical literature exists regarding how to measure performance. Coelli, Estache, Perelman and Trujillo (2003) present a survey of different methodologies to measure performance. The methodologies considered in this study are performance indicators, total factor productivity indexes, and frontiers.

A limitation associated with studies of Central America is the scarcity of data related to the water sector. The first steps of the study involved examining existing data and defining a set of variables to be

collected. During the data collection process, some factors were found to be limiting and others were critical for the success of the process. The subsequent steps of the study related to performance measurement. With key input, output, and quality information, basic performance comparisons can be made. A set of performance indicators commonly used among practitioners in the water sector was calculated to provide a very simple picture of the sector’s performance characteristics in the region. Some of these performance indicators were compared to those presented by the Association of Water and Sanitation Regulatory Entities from Latin America (Asociación de Entes Reguladores de Agua Potable y Saneamiento de las Americas, ADERASA; Corton and Molinary 2008) benchmarking task force in its most recent annual report for Latin American countries.

The availability of data from 2002 to 2005 allowed assessing performance in the region through the calculation of total factor productivity indexes. Finally, to provide a more specific picture of the efficiencies associated to production practices in the region, a production frontier using data envelopment analysis and a stochastic cost frontier were included in the analysis. Differences on some of the values obtained from this variety of performance measures are based on the different assumptions underlying each methodology. A best performer was found consistently through the performance methodologies calculation.

The study is organized as follows. In section two, the data collection process is described. Section three provides an analysis of core performance indicators. Section four utilizes total productivity indexes to examine productivity. Section five and six contain a non parametric and parametric frontier analysis. Section seven concludes.

2. Data Collection Process

The starting point of this study was to build a verifiable data base taking into account the data already available. To this end, the author requested and collected data from ADERASA and the International Benchmarking Network for Water and Sanitation Utilities (IBNET). The ADERASA data base is comprised of data that comes from the regulatory agencies in each Latin American country. This
information is reported by some of the utilities in each sector but not all. Appropriate contact was
established with these utilities to verify the existing data and to obtain missing values. In addition,
Guatemala and El Salvador are not members of ADERASA, so data for these countries were collected for
a first time. The adopted strategy for the collection process was incremental in the sense that data were
sent to the source several times for verification. A new and refined data set for the water sector in Central
America emerged from this process. Nevertheless, only a subset of variables was used for the analysis
because not all countries reported all variables or all years. Consequently, the number of observations was
reduced to allow the data set to be comparable for all utilities and to include all countries.

Several factors were identified as affecting data availability within the region: the on going water
sector restructuring, the low level of water infrastructure in place, and the low presence of information
technology among the service providers.

From an institutional point of view, Costa Rica, Panama and Honduras have independent
regulatory agencies but El Salvador, Guatemala and Nicaragua still have central government bodies
overseeing the water sector. Some of these central government institutions are undergoing restructuring
such as the transfer of sector responsibilities to municipalities. Some countries have just finished a major
restructuring of this type. Because these changes generally imply changes within the company staff, the
flow and registering of data may get interrupted, affecting data collection procedures.

El Salvador, Honduras and Nicaragua show a low level of infrastructure in place. This promotes
the presence of a large number of local independent water providers which limit the data collection
process. Solo (1998) provides a detailed description on the role of these independent providers within the
water sector of Latin American countries.

Overall, the water sector in this region is fragmented given the decentralization of service
provision into municipalities. For instance, for Honduras, with 271 municipalities, and for Guatemala
with more than 300 municipalities the lack of data is evident. Consequently, available data came from the
municipalities serving the largest cities. This fragmentation greatly hampers performance analysis for the
water sector in this region.
Finally, the development of information technology is central to any data collection initiative. According to the United Nations statistics on measuring Information and Communication Technologies, the diffusion index which includes connectivity and access to computers for these countries is approximately 40% for Costa Rica; 30% for Guatemala; 25% for Panama and Honduras and 15% for El Salvador and Nicaragua. Information technology is the core to any structured data collection procedure. The availability of an information system specific for the sector is crucial for any data collection process within this region. Initiatives in this respect are only incipient. In 2004, a workshop hosted by Peru with participation of several Latin American countries representatives from the water sector gathered initial ideas and directed some efforts into the development of a water sector information system common to the region. A similar initiative hosted by El Salvador developed in late 2006.

The presence of technology is necessary but not sufficient for improved information on water utility performance. When designing rules for the sector, the government needs not only to consider the utilities, the main responsible for collecting appropriate data, but also it needs to consider the role that each stakeholder plays in the flow of data within the sector. For instance, the reporting of data to the regulator from the utilities needs to be stated by law and not taken as an informal relationship between the parties. In the same way it is important to establish formal communication channels among all the sector stakeholders, such as environmental or municipal development agencies, in a way that data collection programs and possible data repositories are well identified and efforts are not duplicated.

Costa Rica, Guatemala and Honduras are represented in the analysis by two service providers of different sizes, which are going to be referred as the small and large providers for each of these countries. The rest of countries are represented only by one operator.

3. Sector Performance Indicators

The simplest types of performance indicators are unit-dimensional such as labor productivity, service coverage, and non-revenue water. These indicators focus on a specific area of performance within the production process. Considering the most commonly used performance indicators in this sector, they
have been classified in this study as operational, financial and quality indicators. Since they are benchmarked against those indicators calculated by the ADERASA Benchmarking group, their definitions are kept the same to maintain consistency across the region.

In general, the extent of water service provision can be measured by volume of water, number of connections and population served. These factors provide a good picture about the size of the company providing service. Overall, in the Central American countries the water utilities can be broadly classified as small, medium and large. Table 1 shows volume of water, number of connections and population served in 2005, as well as the values for the performance indicators examined in the following sections.

3.1 Operational Indicators

Water lost or non-revenue water reflects deficiencies in either operational or commercial practices. The extent of water losses may reflect a cost tradeoff between increasing water production and repairing network leaks to keep up with water demand. In other words, to satisfy demand, managers may find it more costly to repair leaks and to control water losses than increase water production. Pipe leaks on the transmission segment require costly maintenance outlays, particularly on long or dispersed networks. Operational water losses arise in transit while in the transport or distribution network, and are calculated as volume of water produced less water delivered to the distribution network. Approximately 5% of water is lost during treatment as material is flushed out, however the starting point distinction between water produced after treatment or water taken by the plant is not considered in this study.

Referring to the distribution system, water losses may be either due to water theft, representing commercial losses, or to pipes’ leaks. It is plausible to argue that given the characteristics of this sector it may be hard for firms to control commercial losses if that entails denying the service to the poorest segments of the population. For the distribution network, water losses are measured as the difference between water delivered to its starting point and water billed. Another way of viewing this indicator is to calculate the ratio of water billed to water delivered to the distribution network which is referred by the ADERASA benchmarking group as an indicator for commercial efficiency. For utilities in the Central
America sample, the median value for this indicator is 55% which is higher than the ADERASA value of 40%.

The metering indicator is calculated as the ratio of the number of connections with an installed meter to number of total connections. Meter acquisition and installation costs are high. In some countries there is a direct allocation of metering costs to the consumer, which generally translates into higher tariffs and/or connection fees. The higher is the level of metering, the higher the possibility of identifying water losses, the more accurate will collection information be and as a consequence, revenues may be higher. Overall, metering median value is 56%, which is lower than the 75% median value for ADERASA members.

Service coverage is calculated as the ratio of population with water service to total population in the area of service of the utility. The median value for water service coverage in this region is 90% which is close to the ADERASA value of 89%. Coverage is equal to 92% for large firms, 66% for medium firms and 85% for the small utilities group.

Water companies with a similar scale, measured by number of connected properties, may have different costs due to differences in network characteristics, such as length. Larger firms could have lower costs due to a large amount of customers per kilometer of pipe, rather than due to scale economies originated from total output. To explore this issue, network density, measured as the ratio of number of connections to network length is the performance indicator considered in this analysis. The median value for network density equals 95 for utilities in Central America. Larger firms have denser networks than medium and small firms. The low coverage and high network density values found for the Guatemala utility suggest that the system may be expanded by increasing the length of network to reach out underserved populated areas. The low coverage and low network density found in Honduras may indicate that the system can be expanded by adding more connections to satisfy underserved population in the area.

The ADERASA benchmarking group utilizes the ratio of volume of water billed to population with water service as an indicator for water consumption. The median consumption value for the region equals 219 liters per person per day, which is slightly higher than the ADERASA value of 172. The group of
smaller companies is characterized by a higher consumption level of 323 liters per person per day as opposed to a lower value of 222 satisfied by larger firms. Figure 2-6 depicts this indicator.

The number of workers per one thousand connections is used in the water sector literature as signaling labor efficiencies or inefficiencies. A large value suggests the company is using a higher than efficient number of workers on its production process. The median value for this indicator equals 6.6, which is twice the value found for ADERASA members suggesting high labor inefficiencies. Note that the number of workers considered for this indicator is a total figure which includes contracting or outsourcing labor.

The availability of several years of data allows us to analyze the changes occurring in number of connections and network length which imply system expansion at different stages of the investment cycle. Clearly, national priorities and funding sources affect the pace and pattern of system expansion.

Figure 1 shows changes in number of water connections and network length to illustrate system expansion as shares of total system expansion for each country. The service provider from Guatemala shows a higher increase in network length with respect to that of number of connections. This may suggest a system expansion of the transportation segment where pipes but not connections are added as opposed to an expansion of the distribution segment where pipes and connections are both added. Alternatively it may indicate earlier stages of the distribution network expansion where customers have not been connected yet. The service provider from Nicaragua presents an opposite situation. Here the increase in number of connections is higher than the increase of length of network. This may be explained by connections added to satisfy commercial and industrial customers who generally do not add to population served.

### 3.2 Financial Indicators

Operating costs for this region include labor and energy costs, chemicals, administrative and sales expenses. Depreciation and finance expenses are considered to be part of total costs. On average, operating costs are $91/connection. Higher values of network density are associated with lower values
for operating costs per connection. The median operating cost per cubic meter is $0.10, half the cost of ADERASA member countries. For the large group of the utilities in the region, the median administrative expense per connection equals $27, whereas it is $34 for the small group. Both values are lower than the similar indicator for ADERASA members which equal $47.

Changes in operating costs for 2002-2005 are of small magnitude. Costa Rica displays a significant increase in cost of workers (54%) and administrative expenses (51%) which may explain a higher increase on its operating costs with respect to that of other utilities. On the other hand, Panama’s increase in operating costs (18%) may be explained by an increase in energy costs (54%). El Salvador displays a decrease in administrative expenses (18%) which may explain the decrease of its operating costs (10%). Service providers which presented an increase in energy costs may be reflecting a combination of increases in input prices and greater utilization of energy inputs to service larger systems. Significant increases in cost of workers could be due to an increased focus on hiring professionals with managerial skills.

### 3.3 Quality Indicators

Compliance with water quality standards has a median value of 95.96% slightly higher than the similar indicator for the ADERASA members. Continuity measured as the number of hours with water service ranges from 20 to 24 hours. Number of complaints per connection (median value) is similar for both ADERASA and Central American utilities (0.21). The median number of leaks per km of pipe is 2.53 for ADERASA members, almost half the value found on Central America countries, 5.19. This suggests a lower degree of pipes service maintenance for Central America water networks compared with the Latin American set of water networks. This also explains the higher value of water lost in the Central American water utilities compared to this value for utilities in Latin America.
4. Total Factor Productivity Analysis

Productivity indexes represent a way of measuring performance over time since they identify production differences between two time periods. A comprehensive exposition and analysis of such measures is found in Balk (2003). The basic idea behind a Total Factor Productivity (TFP) index is to determine how much output is produced due to each unit of input. In practice, there may be more than one output produced from a combination of inputs. In such a case, a TFP index is constructed as the ratio of an output index to an input index. The input index should reflect the relative importance of each input in producing the output(s) and the output index should reflect the relative importance of each output. These relationships are represented by weights. Assume $r$ is the weight given to $k$ outputs and $s$ the weight to $n$ inputs. Equation 1 is the general formula to calculate a TFP index measuring productivity change between two periods of time, say from year 0 to year 1.

$$
\frac{TFP_1}{TFP_0} = \frac{\sum_{k} r_k Y_{k1} / \sum_{n} s_n X_{n1}}{\sum_{k} r_k Y_{k0} / \sum_{n} s_n X_{n0}}
$$

(1)

Three aspects are relevant when setting the above mentioned weights: the selection of the elements that represent the weights, the mathematical or functional form that combine them and whether the weights are the same for the two analyzed periods or not. The weights are generally prices for input factors and costs for outputs. In Equation 1, the weights are assumed the same for both periods but they may not be which yields different alternatives for calculating TFP indexes. When considering the initial set of prices, Equation 1 represents a Laspeyres index while using the final period prices produces a Paasche index. The Fisher index utilizes the geometric mean of the two periods.

Regarding the functional form to relate the weights, the indexes described so far imply a linear functional form. The Tornqvist index uses a logarithm form which is more flexible in reflecting a production technology’s characteristics.
Defining the weight of outputs for calculating these indexes requires detailed information on the production technology of these companies as well more specific data. In order to simplify the analysis such that weights for the outputs are not needed, two sets of TFP measures are calculated: one considering volume of water billed as the output and another considering number of connections. As in most empirical studies related to the water sector, labor and energy are the input factors under consideration. However, not all service providers reported energy volume which limited the calculation of these indexes only to a subset of utilities. The weights for the input factors are calculated as the ratio of their respective costs relative to operating costs. See Table 2 for the results.

The Laspeyeres, Paasche and Fisher TFP indexes yield similar results. This may be explained by the fact that the length of period is short which produces only a small variation when calculating the weights for the different indexes. All companies are more productive from the view point of number of connections as opposed to volume of water billed. Panama is the only country displaying increased productivity over the period when considering both number of connections and volume of water. The productivity increase for Panama ranges from 31% to 53%. Nicaragua displays a very small decrease in productivity ranging from 2% to 5%. Finally, El Salvador depicts a small productivity decrease ranging from 6% to 17%. Presumably, increases in TFP should track decreases in average cost if all the other factors of production besides labor and energy remain constant. The service provider for Panama is expected to perform better with respect to other providers in the frontier performance assessment.

4. Data Envelopment Analysis (DEA)

This section examines each firm’s relative technical efficiency in 2005 using Data Envelopment Analysis (DEA). As in the case of TFP indexes, DEA assumes that the data contains no measurement errors. A difference with respect to the TFP analysis is that the DEA methodology allows us to consider a linear combination of outputs and inputs for the production process without presetting their weights. Rather, these weights are calculated with respect to the combination of these factors found on best producers.
Data Envelopment Analysis (DEA) is the most commonly used non-parametric frontier methodology. It is non-parametric because it does not require the analyst to specify a functional form for the production technology. Charnes, Cooper and Rhodes (1978) were first to present the concept of the relative ranking of decision making units according to their efficiency. A DEA calculation determines simple relationships among variables. For example, utilities that produce far less output than other utilities, which are using the same input levels, are deemed to be relatively inefficient. This methodology is viewed as an “extreme point” method because it compares production of each firm with the “best” producers. Efficiency results from a DEA frontier are contingent to three main factors:

- The composition (homogeneity) of the sample set of firms to be analyzed which is critical in determining the set of best producers to be compared with each firm
- The set of selected inputs and outputs which establishes the comparison terms
- The quality of the data since this methodology assumes that there are no errors

A DEA analysis consists of measuring the efficiency of any firm as obtained by the ratio of weighted outputs to weighted inputs subject to the condition that similar ratios for every firm are less or equal to unity. This relationship is expressed mathematically in Equation 2.

\[
\begin{align*}
Max \ldots \lambda_0 = & \frac{\sum_{r=1}^{s} \alpha_r Y_{r0}}{\sum_{i=1}^{m} \beta_i X_{i0}} \quad S.T. \quad \frac{\sum_{r=1}^{s} \alpha_r Y_{rj}}{\sum_{i=1}^{m} \beta_i X_{ij}} \leq 1 \\
\alpha_r, \beta_i & \geq 0; \quad j = 1 \ldots n; \quad i = 1 \ldots m; \quad r = 1 \ldots s
\end{align*}
\]

Here, \(Y_{00}, X_{0}\) are observed output and input variable vectors of the firm under evaluation; \(\alpha\) and \(\beta\) are the weights to be applied to all units; \(i\) represents an input within a set of \(m\), \(r\) an output within a set of \(s\), and \(j\) one of the \(n\) firms. With this specification, each firm can weight inputs and outputs differently as long as the ratio of their linear combination is less or equal to one. However, this flexibility may raise concerns referring to whether the efficiency value obtained from calculation is a product of efficiency per
se or is originated by the choice of weights. Yet, if the firm turns out to be inefficient even under the most favorable weights, then there is no doubt about the result. The efficiency of firm zero is rated relative to all firms. The maximization gives firm zero the most favorable weight allowed by the constraints.

The output from a DEA exercise is the proportion by which the observed inputs could be contracted if the firm where to operate efficiently. Intuitively this means that the same level of output can be produced optimally with fewer inputs, so it is referred as an input efficiency approach. The implicit assumption is that managers minimize input usage given output level. In the economic literature, this is referred as the Farrell’s measure of input efficiency (Farrell 1957).

Alternatively, the output efficiency approach considers the maximal proportional output expansion with the input vector held fixed, so managers maximize output given a set of inputs. Unless constant returns to scale are assumed, these measures yield different scalar values meaning different distances from the frontier. A comparison of results from input and output oriented models is found in Orea, Roibas and Wall (2004). A study by Banker, Cooper, Seiford, Thrall and Zhu (2004) addresses details and applications of input and output efficiency approaches under variable and constant returns to scale.

For regulated industries, such as the water sector in Central America, an input approach is the natural option given that utilities most generally have service obligations to all customers under a pre-fixed tariff. This approach implies that firms are fully capable of reallocating resources when improving efficiency.

The DEA model critical point is the selection of an appropriate set of inputs and outputs to represent production process with a linear technology. For the Central American water sectors, labor and capital are selected as input factors and volume of water billed and number of connections are the outputs. The amount of energy utilized in the production process was used in the calculation of the TFP indexes and is generally used as an input factor in the production process of water utilities. However data availability is limited for this set of companies. To include all firms in the calculation of the frontier, length of network instead of volume of energy is considered as input factor. Network length is utilized as a proxy to represent capital in the water sector empirical literature. The rationale for doing this is the high
amount of capital necessary to lay down pipes compared to capital needs for the rest of the production infrastructure. Labor is represented in the model by the amount of total workers.

Even when the inefficiency or efficiency of a service provider may be due to its production process per se, a firm can be favored or hindered by country specific circumstances. Indeed, when considering several countries a major challenge is to appropriately account for each country’s political, social and economic differences. The empirical literature on cross country studies for the water sector is limited. Clarke et al. (2004) utilize GNP per capita when analyzing the impact of private participation in the water sectors of Brazil, Argentina and Bolivia. In the electricity sector, Estache, Rossi and Ruzzier (2004) utilize GNP per capita on a DEA and stochastic frontiers to account for countries’ differences when assessing efficiency for South America’s main electricity distribution companies. Zhang, Parker and Kirkpatrick (2007) utilizes GDP per capita in a stochastic model to assess the impact of privatization, regulation and competition on the performance of utilities in the electricity sector of 36 developing countries over a period of 18 years. A study by Estache and Rossi (2008) considers GDP per capita among a set of covariates to capture country particularities on a difference in differences analysis of the electricity sector across 51 developing countries.

The World Bank utilizes the level of gross national income calculated by the Atlas method to classify economies and set the lending eligibilities. The GNI adds to the GDP the income received from other countries, less similar payments made to other countries. The Atlas method uses a three year average of exchange rates to smooth effects of transitory exchange rate fluctuations. The assumption for this study is that variations of GNI may impact the behavior of firms in a country, so that this variable captures country specific economic circumstances affecting the performance of the water sector provider’s service for the countries under analysis. A way to examine this influence under a DEA framework is to include this variable as an “additional resource”. GNI in current dollars from the World Bank Website is included in the model as an input factor.

Utilities in this region have different sizes so it is appropriate to account for firms’ scale when measuring efficiency. A variable returns to scale approach allows increasing or decreasing efficiency
based on size of the firms. Alternatively, a constant return to scale approach means that firms are able to linearly scale the inputs and outputs without increasing or decreasing efficiency. The ratio of these two measures produces the scale impact on efficiency for each firm. Table 3 shows results for technical efficiency and scale impacts under variable returns to scale.

Overall, higher efficiency values are obtained when including network length and subsequently adding GNI. Including each of these variables implies adding input dimensions to the calculation of the frontier which translates into higher efficiencies as it better describes production characteristics. In addition, the sensitivity of results indicates the appropriateness of considering countries differences when explaining efficiency within this region.

Results from Table 3 indicate that the service providers in Panama, El Salvador, Nicaragua, and the small providers in Costa Rica, Guatemala and Honduras are all 100% efficient. The large provider from Costa Rica could produce the same output with approximately 63% of its resources, in this case, labor, network length and GNI. This means that this provider is only 37% inefficient. In the case of the service provider for Honduras, 44% of its efficiency comes from considering its size in comparison to other providers.

Figure 2 shows a production frontier depicting labor as the input factor and number of connections as output. This graph is incomplete in the sense that is showing only one dimension of the frontier given that the other input and out factors are not shown. Nevertheless, it gives an idea of the firm’s relative positions. For instance, El Salvador, and Nicaragua produce a very similar amount of connections. However the amount of labor differs widely. While El Salvador is found 100% efficient, Nicaragua can work with 60% of its labor amount if considering labor the only input factor. Costa Rica and Nicaragua levels of efficiency are very close (62% and 59% respectively).

In this DEA analysis, only data from 2005 were utilized. The Malmquist technical efficiency change component for the period 2002-2005 is calculated for all firms except for the large service provider from Honduras. A Malmquist index measures the Total Factor Productivity change between two time periods utilizing the ratio of the distances of each data point relative to a common production
technology. When calculating this index it is common practice to utilize DEA to calculate these distances. Following Fare, Grosskopf, Norris and Zhang (1994), and considering the input perspective already selected, a Malmquist technology change component based on the geometric mean of the considered periods is defined as follows:

\[
\frac{TFP_t}{TFP_0} = \frac{D_0(Y_t, X_t)}{D_t(Y_t, X_t)} \left[ \frac{D_t(Y_0, X_0)D_t(Y_t, X_t)}{D_0(Y_0, X_0)D_0(Y_t, X_t)} \right]^{1/2}
\]

The first ratio term indicates a measure of input-oriented technical efficiency change for the analyzed period (the catching up effect or movement towards the frontier). Negative values indicate efficiency has declined over the period (the initial efficiency value is higher than the final value). Positive values indicate increased efficiency. A value equal to zero indicates no efficiency change. The term within brackets represents technical change calculated as the geometric mean of the shift in technology between two periods. It is important to notice that while a TFP index is calculated only on reference to a particular firm – a firm change of productivity over time – the efficiency change component of the Malmquist index is calculated with respect to the movement of a firm towards the optimal frontier, which is determined by a group of firms.

The utilities from El Salvador, Nicaragua, and the small utilities from Costa Rica and Honduras show no change in technical efficiency over the period. The utilities from Panama and Guatemala show an increase in efficiency of 8% and 3% respectively. The large service provider from Costa Rica shows a decrease in efficiency of 2%.

5. Stochastic Frontier

So far, inputs and outputs have been considered using calculations from non-deterministic approaches to examining efficiency and productivity in a production technology framework. This section examines cost efficiency by statistically estimating cost relationships according to input prices, given a level of output produced. The ideal framework would be to completely specify a cost function including outputs, input prices and those specific factors capturing possible cost differences among firms and
countries. The limitation associated with data availability restricts the analysis to the inclusion of only four explanatory variables in the economic model. The merit of performing this econometric exercise is that this methodology recognizes the presence of data errors which is an important element for the Central American analysis.

The DEA approach considers efficiency with respect to the best performers, given the variables selected whereas an estimated cost frontier is a measure of central tendency considering all firms not just those on the frontier. The economic model is a cost function specified by volume of water billed (VolBil), price of labor, price of energy, and GNI. Operational costs plus administrative expenses divided by number of connections comprise operating costs per connection which is the dependent variable. All variables are divided by number of connections, to control for size of the firms. The price of labor (LabPrice) is equal to total cost of workers divided by total number of workers. The price of energy (EnegPrice) is total energy expenses divided by length of the network. The economic model translates into the following specification (time and firm specific subscripts are omitted for simplicity):

\[
\text{UnitOpCost} = \alpha + \varphi \text{GNI} + \gamma \text{VolBil} + \beta_L \text{Lab Price} + \beta_E \text{Eneg Price}
\]  

(4)

Where \( \alpha, \varphi, \gamma, \beta_L, \) and \( \beta_E \) are parameters to be estimated. All variables are in natural logarithm form. The data set is an unbalanced panel covering 2002 to 2005. Empirical researchers often introduce a time trend in the model to capture possible technology shifts explaining technological changes. Given the short period of time plus the limitation on the number of explanatory variables, GNI instead of a time trend is included to capture possible economic shocks occurred over the period.

Nevertheless, time is included in the estimation of the cost frontier to explain possible changes of efficiency over time. Here the basic assumption is that these countries have changed efficiency behavior over time, independently. A stochastic cost panel data frontier using the Battese and Coelli (1993) exponential functional form for the inefficiency term is adopted to estimate the model. The weakness of this specification is that it assumes all firms have followed the same trend, which is a restriction but still a
reasonable assumption for the water utilities within the region. The specification for the econometric model includes Equation 4 plus the error term \( \varepsilon \) specified as follows:

\[
\varepsilon_{it} = u_{it} + v_{it} \quad \text{where} \quad u_{it} = u_{it} \exp[-\eta(t - T)]; u_{it} \text{ iid } \sim N(0, \sigma_u^2) \]

The idiosyncratic error term \( (\nu) \) is independently and identical distributed as \( \sim N(0, \sigma_v^2) \) and independently from regressors. The only panel-specific effect is the random inefficiency term \( (\mu) \). The frontier estimator assumes the effects (error term) are not correlated to the explanatory variables. The statistically insignificant coefficient for the parameter eta implies that inefficiency has not changed over time, or if it has, it does not follow an exponential path. The model seems to provide a good fit and variables have good explanatory power (illustrated by the high statistically significance of most coefficients). According to estimation results, the operating costs for the region are represented by the following equation.

\[
\text{UnitOpCost} = (-5.3) - 0.1*GNI + 0.6*VolBil + 0.7*\text{Lab Price} - (0.1)*\text{Eneg Price} \quad (6)
\]

The signs of coefficients are as expected and the statistical significance is high for the output variable and price of labor. Energy price and GNI are not statistically significant. The coefficients’ magnitude for the statistically significant variables is high, indicating a high economic impact on costs.

An increase of 1% in volume of water per connection produces an increase in unit costs of approximately 0.6%. An increase in price per worker per connection of 1% produces an increase in unit costs of approximately 0.7%. The impact on costs from increases in price of labor is the largest relative to the impacts of other variables. This finding reinforces the importance of labor costs within the sector.

Increasing volume has a less than proportional increase effect on costs. On average it indicates the presence of economies of scale in the region. Table 4 shows country ranks according to how far each firm is from the efficient cost frontier and possible reduction of operating costs by 2005. This reduction is calculated as the ratio of the estimated inefficiency to actual operating costs.

The differences between results from the DEA frontier and the estimated cost frontier can be explained by the fact that when calculating the DEA frontier we were looking at a contraction of inputs
for a given level of output. A cost frontier looks at minimizing costs given input prices and output. When assuming the minimum set of inputs for a given level of output, the DEA abstracts from other factors influencing the production process, such as the price of inputs. However, Panama is been a constant high performer for all methodologies. Results for the other countries may be considered on a case by case basis.

7. Concluding Observations

A major contribution from this study is the creation of a unique data base and a comprehensive data collection process. The data collected is very representative from a sector wide perspective, considering total water service coverage in the region. The quality of the data set is considered good in the sense that it came from and it was reviewed and certified by data owners.

A major conclusion from this analysis points towards additional efforts for improving data collection procedures in the region. Besides the scarce presence of information technology limiting record keeping within these utilities, difficulties also may be due to the fragmented service provision in some countries. A higher level of coordination is needed if data are to be collected and trends analyzed. Such an initiative may require an analysis of stakeholders’ responsibilities regarding monitoring and storage of data. Coordination is needed among stakeholders regarding what and how to collect data. Information technology is central to any structured data collection procedure: the availability of an information system specific for the sector is crucial for any data collection process within this region. Government, policy makers and fund providers need to consider the role of technology improvement in the region.

In the process of identifying segments of the industry with no data, policy-makers, regulators and managerial staff have been encouraged to expand efforts to seek disaggregated data. Such data are necessary for further quantitative analysis, providing more complete information regarding sector performance.

A set of methodologies was utilized to assess sector performance, such as performance indicators, total factor productivity analysis, production and cost frontiers. According with the results the study did
identified different stages of the investment cycles for the group of utilities. On their water system expansion, Guatemala seems to be adding more pipes and Nicaragua more connections. While the former is trying to reach under-served areas, the latter is trying to satisfy local customers demand. Investment for these two types of systems expansion is different. Adding more pipes require higher levels of capital than adding only connections.

The amount of non revenue water is higher in Central American countries (55%) with respect to that in Latin American countries (40%). This may be a consequence of poor metering systems as it is reflected in a 56% metering value for these countries with respect to a 75% for the sample of ADERASA Latin American countries. It may also be explained by the amount of pipe leaks which is very high when compared to the same indicator for Latin American countries.

Large firms have denser networks than smaller firms, reflecting the low investment capacity of small providers. Guatemala shows a low level of coverage compared to its high network density which is consistent with the extension of its network system through adding more pipes as it was previously mentioned. In Honduras, the low coverage ratio and network density implies that the expansion of the system should be through adding more connections to satisfy under-served population in the area. Water consumption is higher in Central American countries when compared to Latin American countries which reflect abundant sources of water in the region. Labor inefficiencies are twice those found in Latin American Countries.

On the financial side operating costs per cubic meter of water are half the one found in Latin American countries. Administrative expenses are found to be higher in small firms. Changes in operating costs over the period are reflecting changes in energy costs, as in the case of Panama, or labor costs as the case of Costa Rica.

When trying to assess changes over the period, calculating different total factor productivity indices yield similar results given the short period analyzed. Panama is the only country displaying an increase in productivity when considering labor and energy as input factors over the period 2002-2005.
This country also shows an increase in efficiency as a result of calculating the Malmquist efficiency change component. Nicaragua is the only country with a small decrease on efficiency change (2%).

When calculating a DEA technical frontier for 2005, higher efficiency values are obtained from models that include GNI as an additional factor in addition to number of workers and network length. This indicates the importance of including country characteristics in explaining efficiency within the region.

Results indicate that the service providers in Panama, El Salvador, Nicaragua, and the small providers in Costa Rica, Guatemala and Honduras are all 100% efficient with respect to the whole group. Part of this efficiency for El Salvador, Nicaragua and the small providers is due to their scale. The fact that such a large number of firms are found to be 100% efficient indicates that the group of firms is not heterogeneous enough in the sense that the set of input and output factors considered are not sufficient to explain possible production differences among these countries.

Regarding the stochastic cost frontier estimation, results indicate that the highest impact on operating costs per connection comes from labor prices, when considering input prices and output. Estimated firm’s ranks according to how far is each firm from the cost frontier shows Nicaragua as the most efficient firm. The service provider from Panama is ranking number 2 which coincides with the fact that its service provider is positioned as the best performer from other methodologies. The ranking of other providers need to be assessed on a case by case basis since some results are contradictory. Data limitations affect model specification: not all elements affecting cost are in the model.

For regulatory agencies, related government institutions and funding agencies, this study may contain additional information for their strategic planning and decision making processes. This study should be viewed as a first step in the analysis of water utilities in Central America. As additional years become available and more utilities provide information, analysts will be able to conduct much more thorough analyses of sector performance. The results presented here are
expected to serve as a catalyst for more comprehensive data collection/verification initiatives in the region and for additional quantitative studies.

Table 1. Summary of operational performance indicators by 2005

<table>
<thead>
<tr>
<th>Country/Units</th>
<th>Vol Del</th>
<th>Vol Lost</th>
<th>Num Conn</th>
<th>Met Del</th>
<th>Met Lost</th>
<th>Vol Serv</th>
<th>Pop Cov</th>
<th>Net Length</th>
<th>Net Dens</th>
<th>work/1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panama</td>
<td>452</td>
<td>58</td>
<td>448</td>
<td>41</td>
<td>126</td>
<td>2,004</td>
<td>92</td>
<td>4,727</td>
<td>95</td>
<td>5.6</td>
</tr>
<tr>
<td>Costa Rica-L</td>
<td>305</td>
<td>49</td>
<td>457</td>
<td>94</td>
<td>76</td>
<td>1,978</td>
<td>99</td>
<td>6,437</td>
<td>71</td>
<td>6.7</td>
</tr>
<tr>
<td>El Salvador</td>
<td>259</td>
<td>-</td>
<td>619</td>
<td>55</td>
<td>84</td>
<td>3,093</td>
<td>90</td>
<td>4,391</td>
<td>141</td>
<td>4.2</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>257</td>
<td>44</td>
<td>457</td>
<td>48</td>
<td>39</td>
<td>2,870</td>
<td>91</td>
<td>4,604</td>
<td>99</td>
<td>6.7</td>
</tr>
<tr>
<td>Guatemala-L</td>
<td>122</td>
<td>55</td>
<td>195</td>
<td>84</td>
<td>81</td>
<td>1,045</td>
<td>93</td>
<td>5,013</td>
<td>128</td>
<td>7.5</td>
</tr>
<tr>
<td>Honduras-L</td>
<td>75</td>
<td>63</td>
<td>105</td>
<td>35</td>
<td>67</td>
<td>707</td>
<td>69</td>
<td>2,800</td>
<td>38</td>
<td>10.5</td>
</tr>
<tr>
<td>Costa Rica-S</td>
<td>28</td>
<td>55</td>
<td>48</td>
<td>97</td>
<td>66</td>
<td>228</td>
<td>100</td>
<td>678</td>
<td>71</td>
<td>2.5</td>
</tr>
<tr>
<td>Honduras-S</td>
<td>10</td>
<td>44</td>
<td>11</td>
<td>77</td>
<td>157</td>
<td>53</td>
<td>71</td>
<td>77</td>
<td>144</td>
<td>6.6</td>
</tr>
<tr>
<td>Guatemala-S</td>
<td>7.6</td>
<td>-</td>
<td>10</td>
<td>56</td>
<td>183</td>
<td>42</td>
<td>72</td>
<td>232</td>
<td>42</td>
<td>9.6</td>
</tr>
</tbody>
</table>

![WaterConx vs NetLeng](image.png)

Figure 1. Changes in number of water connections and network length
### Table 2. Total factor productivity indexes for 2002-2005

<table>
<thead>
<tr>
<th>Index</th>
<th>Output Variable</th>
<th>Laspeyres, Volume</th>
<th>#Connect</th>
<th>Paasche &amp;Fisher Volume</th>
<th>#Connect</th>
<th>TFP</th>
<th>Tornqvist #Connect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panama</td>
<td></td>
<td>51%</td>
<td>53%</td>
<td></td>
<td></td>
<td>31%</td>
<td>32%</td>
</tr>
<tr>
<td>El Salvador</td>
<td></td>
<td>-17%</td>
<td>-6%</td>
<td></td>
<td></td>
<td>-9%</td>
<td>0</td>
</tr>
<tr>
<td>Nicaragua</td>
<td></td>
<td>-5%</td>
<td>1%</td>
<td></td>
<td></td>
<td>-5%</td>
<td>-2%</td>
</tr>
</tbody>
</table>

### Table 3. DEA technical efficiency and scale impact on efficiency for 2005

<table>
<thead>
<tr>
<th>Inputs</th>
<th>Variable Returns</th>
<th>Scale Impact on Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Labor Netlength</td>
<td>Labor Netlength GNI2005</td>
</tr>
<tr>
<td>Panama</td>
<td>1.00 1.00</td>
<td>1.00 1.00</td>
</tr>
<tr>
<td>El Salvador</td>
<td>1.00 1.00</td>
<td>1.00 0.73</td>
</tr>
<tr>
<td>Costa Rica-L</td>
<td>0.62 0.62</td>
<td>0.63 0.60</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>0.59 0.70</td>
<td>1.00 0.59</td>
</tr>
<tr>
<td>Guatemala-L</td>
<td>0.52 0.85</td>
<td>0.85 0.72</td>
</tr>
<tr>
<td>Honduras</td>
<td>0.34 0.38</td>
<td>0.99 0.99</td>
</tr>
<tr>
<td>Costa Rica-S</td>
<td>1.00 1.00</td>
<td>1.00 1.00</td>
</tr>
<tr>
<td>Guatemala-S</td>
<td>1.00 1.00</td>
<td>1.00 0.62</td>
</tr>
<tr>
<td>Honduras-S</td>
<td>1.00 1.00</td>
<td>1.00 0.39</td>
</tr>
</tbody>
</table>

23
Figure 2. Constant returns to scale frontier

Table 4. Efficiency rank and possibilities of operating costs reductions by 2005

<table>
<thead>
<tr>
<th>Country</th>
<th>Frontier Rank</th>
<th>Possible Cost Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nicaragua</td>
<td>1</td>
<td>2%</td>
</tr>
<tr>
<td>Panama</td>
<td>2</td>
<td>2%</td>
</tr>
<tr>
<td>El Salvador</td>
<td>3</td>
<td>3%</td>
</tr>
<tr>
<td>Costa Rica-S</td>
<td>4</td>
<td>3%</td>
</tr>
<tr>
<td>Guatemala-S</td>
<td>5</td>
<td>4%</td>
</tr>
<tr>
<td>Costa Rica-L</td>
<td>6</td>
<td>3%</td>
</tr>
<tr>
<td>Guatemala-L</td>
<td>7</td>
<td>12%</td>
</tr>
<tr>
<td>Honduras-L</td>
<td>8</td>
<td>14%</td>
</tr>
</tbody>
</table>

ACKNOWLEDGEMENTS
This study is part of a Public Utility Research Center project funded by the Inter-American Development Bank, Natural Resources and Environment Division during 2007. [RS-T1271: Benchmarking de Empresas Públicas de Agua y Saneamiento en Centroamérica].
LIST OF REFERENCES


