Received 21 January 2014
Revised 3 October 2014
Accepted 17 October 2014

Benchmarking the efficiencies of Indonesia's municipal water utilities using Stackelberg data envelopment analysis

Andreas Wibowo
Research Institute for Human Settlements, Ministry of Public Works, Bandung, Indonesia, and
Hans Wilhelm Alfen
Chair of Construction Economics, Bauhaus University, Weimar, Germany

Abstract

Purpose – The purpose of this paper is to present a yardstick efficiency comparison of 269 Indonesian municipal water utilities (MWUs) and measures the impact of exogenous environmental variables on efficiency scores.

Design/methodology/approach – Two-stage Stackelberg leader-follower data envelopment analysis (DEA) and artificial neural networks (ANN) were employed.

Findings – Given that serviceability was treated as the leader and profitability as the follower, the first and second stage DEA scores were 55 and 32 percent (0 percent = totally inefficient, 100 percent = perfectly efficient), respectively. This indicates sizeable opportunities for improvement, with 39 percent of the total sample facing serious problems in both first- and second-stage efficiencies. When profitability instead leads serviceability, this results in more decreased efficiency. The size of the population served was the most important exogenous environmental variable affecting DEA efficiency scores in both the first and second stages.

Research limitations/implications – The present study was limited by the overly restrictive assumption that all MWUs operate at a constant-return-to-scale.

Practical implications – These research findings will enable better management of the MWUs in question, allowing their current level of performance to be objectively compared with that of their peers, both in terms of scale and area of operation. These findings will also help the government prioritize assistance measures for MWUs that are suffering from acute performance gaps, and to devise a strategic national plan to revitalize Indonesia's water sector.

Originality/value – This paper enriches the body of knowledge by filling in knowledge gaps relating to benchmarking in Indonesia's water industry, as well as in the application of ensemble two-stage DEA and ANN, which are still rare in the literature.

Keywords Benchmarking, Data envelopment analysis, Indonesia, Artificial neural network, Municipal water utility, Serviceability, Profitability

Paper type Case study

1. Introduction

In their effort to meet Millennium Development Goals, the Government of Indonesia (GoI) set the national target for piped water service coverage at 68.87 percent of Indonesia's

The first author is indebted to the Alexander von Humboldt Foundation for providing a postdoctoral fellowship under the Georg-Forster for Experienced Researchers program. This paper was completed during the first author's third research stay at the Chair of Construction Economics, Bauhaus University in Weimar. The authors would also like to express their gratitude to three anonymous reviewers for their useful comments and suggestions that improve the quality of this paper.

Benchmarking: An International Journal
Vol. 22 No. 4, 2015
pp. 588-609

1463-5771
DOI 10.1108/BIJ-01-2014-0009

total population by 2015 (78.19 percent of the urban population and 61.60 percent of the rural population; Ministry of Public Works, 2010). However, while the GoI did make considerable progress in this area – urban and rural coverage in 2009 were 47.23 and 11.5 percent, respectively, indicating increases from the 2004 levels of 41 and 8 percent, respectively – it fell short of its initial targets by about 13 and 18.5 percent, respectively.

One of the critical success factors in meeting the MDGs is the performance of municipal water utilities (MWUs), which dominate water operations, accounting for more than 90 percent of production. There have been large- and small-scale cooperations between MWUs and private investors, following the public-private partnership (PPP) and business-to-business (B-B) models; however, these cooperations only represent a small fraction of water service providers.

In 2012, out of 328 MWUs evaluated in four areas of fitness – financial, quality of services, operational, and human resources – only about half (171) were classified as "fit" while the rest were classified as "less fit" (101) or "unfit" (56; Supporting Agency for the Development of Drinking Water Provision System, 2013). If the poor performance of MWUs has not been resolved, it appears unlikely that the GoI will be able to reach the MDG targets by 2015; further, by extension, it would be unlikely that they would then attain their target of 100 percent of the population having access to drinking water by 2019.

The poor performance of public water utilities often stems from inefficiencies and mispricing (Hassanein and Khalifa, 2007; Janakanarajan *et al.*, 2006; Tan, 2012). Those inefficiencies need to be measured systematically to decide upon appropriate remedial actions, and benchmarking is a robust measurement tool for this purpose. This multi-faceted technique can be used to identify operational and strategic gaps and to search for best practices that would eliminate such gaps (Yasin, 2002).

To the best of our knowledge, no methodical efforts have yet been made to yardstick measure the efficiencies of Indonesia's MWUs. Although the existing performance evaluation conducted by the GoI provides an initial rough assessment, it does not sufficiently address the questions of which MWUs operate most efficiently, and in which areas inefficient MWUs perform the most poorly. As the input data from this GOI assessment were ordinally scaled, the performance evaluation also did not measure to what extent the benchmarking gaps may have occurred.

In the present paper we measured relative efficiencies for Indonesia's MWUs, benchmarked against the best performers. We also measured the impact of exogenous environmental variables on these efficiencies. A two-stage data envelopment analysis (DEA) using the Stackelberg leader-follower model was employed to benchmark efficiencies, and artificial neural networks (ANNs) were used to investigate the role of exogenous environmental variables.

The present paper makes a number of practical and theoretical contributions. First, the research findings will enable the management of the MWUs in question to objectively compare their current level of performance with that of their peers, both in terms of scale and area of operations; this will allow individual MWUs to understand their relative position and use this knowledge as a tool for self-improvement. Second, the findings will facilitate the GoI's ability to prioritize assistance measures for MWUs suffering from acute performance gaps, and may help the GoI to devise a strategic national plan to revitalize the country's water sector. Finally, the present paper enriches the body of knowledge by filling in knowledge gaps relating to benchmarking in Indonesia's water industry, as well as in the application of ensemble two-stage DEA and ANNs, which are still rare in the literature.

2. Indonesia's water supply sector

The history of Indonesian MWUs may date back to as early as the Dutch colonialism of the nineteenth century; however, the real precursor of contemporary MWUs was established in 1962, when the GoI enacted Act No. 5, which regulated local enterprises, including those responsible for waterwork systems. The responsibility for providing water, from production to distribution, traditionally rested with the central government; with the establishment of a number of waterwork agencies, some functions were delegated and the central government was only responsible for water treatment plants. However, this sharing of responsibilities did not always work smoothly, due to agencies' limited resources.

The roles of waterwork agencies began to take shape in the early 1970s, as the GoI embarked upon massive infrastructure development programs. These plans included water supply projects in hundreds of cities following the oil boom that allowed Indonesia to enjoy windfall profits from skyrocketing oil prices on the international market. In 1975, the central government, via the Ministry of Internal Affairs, released Instruction No. 26, which changed the status of waterwork agencies into local water enterprises (*Perusahaan Daerah Air Minum*) under the ownership of local governments; this has remained the format of MWUs to this day. As the sole shareholders, local governments have discretion as to the management of utilities, including appointing top-level management officers and setting water charges, although water charges are also subject to local parliamentary approval.

At present, a total of 483 MWUs have been in operation in 33 provinces. However, this number changes as new districts are formed, a trend that has been especially strong in Indonesia since the 1999 enactment of the Law on Local Governments. As a result of the 2005 Government Regulation No. 16, the role of central government has shifted from operator to regulator, giving MWUs full responsibility regarding the water supply. As with the state electricity utility, MWUs also have the first right of refusal, in the sense that they are offered the opportunity to supply water services to the public before any other corporations. This government regulation also mandated the establishment of a supporting agency for water provision systems, a non-structural institution under the auspices of the Ministry of Public Works. This agency regulates Indonesia's water sector, but has no direct control over MWUs.

3. Benchmarking efficiencies in the water sector

Benchmarking, according to its simplest definition, is the search for industry best practices that lead to superior performance (Camp, 1989). Inherent in this definition are two issues: who is the best and what are best practices? (Bell and Morey, 1994) and whom it is compared against and what is compared? (Barber, 2004). Benchmarking is more formally defined as the continuous process of measuring products, services, and practices against the company's toughest competitors, or against companies recognized as industry leaders (Camp, 1992).

The use of benchmarking techniques has appeared in many academic journals and technical reports, for diverse fields of inquiry (Luu *et al.*, 2008). The water sector is no exception, although the number of studies focussing on efficiency benchmarking in this industry is small.

Love *et al.* (1998) summarized the main learning experiences from three diverse benchmarking projects carried out in the Northwest Water section of the Regulated Utility Division of United Utilities in the UK between 1995 and 1996. In this paper, the researchers demonstrated that benchmarking is both a useful theoretical concept and quality-management technique.

Corton (2003) used data from the Peruvian water sector to illustrate how yardstick comparisons can improve sector performance. The use of a regression model for operating costs allowed for the establishment of a cost frontier. The sampled utilities were ranked according to their efficiency levels using deviations of actual operating cost from the average.

Also using the data sets from the Peruvian water sector, Lin (2005) demonstrated how the introduction of quality variables affects performance comparisons across utilities. Lin also demonstrated how using different specifications of stochastic frontier analysis (SFA; half-normal and exponential models, both with and without quality as an output) can incorporate quality into benchmarking studies.

Braadbaart (2007) tested the hypothesis that benchmarking affects both transparency and economic performance in the public sector, based on 1989-2000 time-series data on benchmarking and non-benchmarking water utilities in the Netherlands. Braadbaart's test confirmed that benchmarking enhances transparency and performance, but these findings did not support the hypothesis that utility managers will only tighten financial discipline when benchmarking is embedded in a system of managed competition.

Hon and Lee (2009) investigated the efficiencies of the Malaysian water industry, using both DEA and regression analysis. They found mean technical efficiency to be about 66 percent, and demonstrated significant differences in efficiencies across different states.

Employing the DEA procedure, Norton and Weber (2009) compared efficiencies among American water utilities; they concluded that public utilities were most efficient overall, followed by private, not-for-profit utilities; private, for-profit utilities were the least efficient. They also inferred that public utilities were more efficient than private, not-for-profit utilities when serving a variety of customer types, while private, not-for-profit utilities were more efficient when serving a single customer type.

Corton and Berg (2009) computed total factor productivity indicators for water utilities in six Central American countries: Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua, and Panama, using two quantitative methodologies, DEA and SFA. These two techniques yielded different results due to differences in underlying assumptions and techniques for comparing performance; DEA produced higher efficiency scores for models that included the gross national income factor, in addition to the number of workers and network length.

Romano and Guerrini (2011) used DEA to measure and compare the efficiencies of Italian water utilities. They found that ownership structure, size, and geographical location all impacted the performance of water companies, although at different degrees of significance. They also found that publically owned utilities were more efficient than utilities with mixed ownership, as well as that economies of scale emerged for medium-sized firms with more than 50,000 customers.

Singh *et al.* (2011) used DEA with an input-oriented variable-returns-to-scale (VRS) assumption to evaluate the efficiencies of 35 North Indian urban water utilities. They affirmed that most of the sampled utilities performed poorly, with significant opportunities for improvement in operation and maintenance (OM) expenditures, staff size, and water losses. However, they acknowledged two key limitations of their study: their use of a small sample and their exclusion of environmental factors.

Running DEA on a large sample data set of Japanese water utilities, Marques *et al.* (2014) found that the average levels of inefficiency were 57 and 24 percent under the constant-returns-to-scale (CRS) model and the VRS model, respectively. They argued that citizens would benefit most from improving sector efficiencies and transferring funds to more innovative sectors, rather than using scarce funds to subsidize water distribution.

Our literature review results suggest that the methods used to benchmark water utility companies can be broadly classified into two groups: parametric and non-parametric methods, with the most widely used techniques being SFA and DEA, respectively. Although SFA models have some advantages over DEA models in terms of their capacity to account for noise and their potential for conventional hypothesis testing, they pose more difficulties in accommodating multiple outputs (Mugisha, 2014). Thus, DEA fits with the objective of the present paper, as the problem entails multiple inputs and outputs. However, the large body of literature that has used DEA as the framework for analysis has typically focussed on technical efficiencies, with the relationships between inputs and outputs modeled as a single stage.

On another front, most MWUs must balance delivering quality services to their customers (serviceability) with remaining profitable enough to ensure that they can sustainably provide water services to the public (profitability). Serviceability and profitability are closely related goals and cannot be dealt with separately, as they are likely to be dependent on one another. In light of this situation, we characterized the internal structure of MWUs as two input-output subsystems, with one system representing serviceability (focussing on technical outputs, and not on inputs) and another representing profitability (focussing on financial outputs only). Previous studies did not sufficiently resolve these certain issues and we identified them as knowledge gaps to fill.

4. Research methodology

In this section we will explain the research methodology we used to meet our objectives. This section mainly consists of model development, a brief overview of how the leader-follower model was applied, identification of input and output measures and exogenous variables, data collection, and data analysis methods.

4.1 Model development

We situated our model within the DEA framework, with MWUs designated as decision-making units (DMUs). DEA has two distinct advantages: first, it is non-parametric, in the sense that a priori specification of the production function is not required; second, and perhaps more importantly, it easily handles multiple inputs and multiple outputs, and allows for direct comparisons of production possibilities without requiring additional input price data (Collier *et al.*, 2011).

An array of alternatives to DEA-based approaches is available to measure the efficiencies of subsystems and overall systems. For example, one simple option would be to run two separate conventional DEA models for each of the two stages; however, this would result in a lack of coordination among the efficiencies of the whole system (Cook *et al.*, 2011). Moreover, Chen and Zhu (2004) have mathematically proven that the standard DEA model can measure efficiency in each stage, but it cannot accommodate a two-stage efficiency with an intermediate measure in a single implementation.

We did not attempt to balance the two subsystems, but rather emphasized the priority of the first subsystem, as it is more important. This implies that the efficiency of the second subsystem will be conditional on that of the first. Our model would then best be represented by a two-stage Stackelberg leader-follower DEA model. Under this model, the efficiencies of the first subsystem (i.e. stage) are treated as the leader and those of the second subsystem as the follower. The principal objective functions of this model are to maximize the leader's efficiency scores first, and then to maximize those of the follower, with the constraint that the multipliers used must be such that the first-stage score remains unchanged (Cook *et al.*, 2011).

Furthermore, under the DEA framework, DMUs are typically assumed to operate within homogenous environments. However, this assumption is no longer valid when environmental performance is influenced by variables beyond management's control (Macpershon *et al.*, 2013). Within a relative-efficiency context, when DMUs operating in different environments are compared, those that operate in less desirable environments are at a disadvantage (Saraiya, 2005). Therefore, the impact of exogenous environmental variables on efficiencies is also an interesting area for investigation. Figure 1 schematically sums up our framework for analysis.

4.2 Stackelberg leader-follower DEA model

Liang *et al.* (2008) formulated the Stackelberg leader-follower DEA model as follows: let us suppose that a set of n DMUs ($j = 1, 2, \ldots, n$) has m inputs x_{ij} ($i = 1, 2, \ldots, m$) in the first stage and D outputs z_{dj} ($d = 1, 2, \ldots, D$) resulting from this stage. These outputs, which are referred to as intermediate measures, become the inputs for the second stage, whose outputs are y_{rj} ($r = 1, 2, \ldots, s$).

Under the CRS and output-oriented assumptions, the first-stage and second-stage efficiencies for DMU_j are defined as follows:

$$e_j^1 = \frac{\sum_{d=1}^{D} w_d z_{dj}}{\sum_{i=1}^{m} v_i x_{ij}} \quad \text{and} \quad e_j^2 = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{d=1}^{D} \hat{w}_d z_{dj}} \tag{1}$$

where e_j^1 is the first-stage efficiencies; e_j^2 the second-stage efficiencies; and w_d, v_i, u_r, $\hat{w}_d$ are unknown non-negative weights.

Assuming an output-oriented model, and assuming that the first stage is the leader, for a specific DMU_o:

$$e_o^{1*} = \max \sum_{d=1}^{D} w_d z_{do}$$

$$\text{s.t.} \sum_{d=1}^{D} w_d z_{dj} - \sum_{i=1}^{m} v_i x_{ij} \leqslant 0, j = 1,\ 2, \ldots,\ n$$

$$\sum_{i=1}^{m} v_i x_{io} = 1$$

$$w_d \geqslant 0,\ d = 1,\ 2,\ \ldots,\ D,\ v_i \geqslant 0,\ i = 1,\ 2,\ \ldots,\ m \tag{2}$$

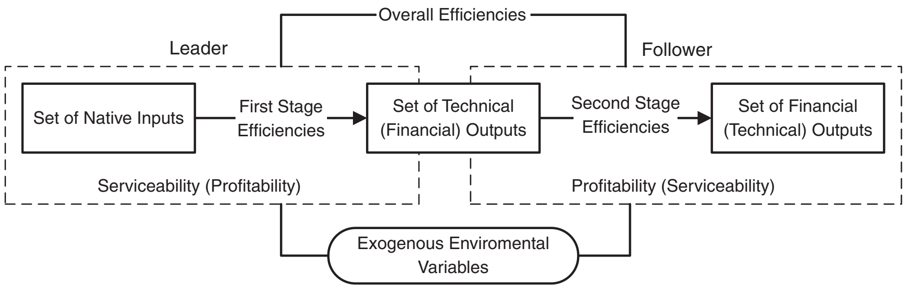

Figure 1. Model development under data envelopment analysis framework

where e_o^{1*} is the conventional DEA score for DMU_o. Setting $w_{d}, = \hat{w}_d$, the second-stage efficiencies for DMU_o can be calculated as follows:

$$e_o^{2*} = \max \sum_{r=1}^{s} u_r y_{ro} / e_o^{1*}$$

$$\text{s.t.} \ \sum_{r=1}^{s} u_r y_{rj} - \sum_{d=1}^{D} w_d z_{dj} \leqslant 0, \ j = 1, \ 2, \ldots, \ n$$

$$\sum_{d=1}^{D} w_d z_{dj} - \sum_{i=1}^{m} v_i x_{ij} \leqslant 0, \ j = 1, \ 2, \ \ldots, \ n$$

$$\sum_{i=1}^{m} v_i x_{io} = 1$$

$$\sum_{d=1}^{D} w_d z_{do} = e_o^{1*}$$

$$w_d \geqslant 0, d = 1, \ 2, \ \ldots, \ D, \ v_i \geqslant 0, \ i = 1, \ 2, \ \ldots, \ m, \ u_r \geqslant 0, \ r = 1, \ 2, \ \ldots, \ s \tag{3}$$

Unlike the two-stage centralized DEA model (Liang *et al.*, 2006), e_o^{1*} and e_o^{2*} are unique, and are not affected by the possibility of multiple solutions (Cook *et al.*, 2011). This optimal solution is another advantage of using the leader-follower model. The efficiencies of the overall two-stage process are a product of individual efficiency measures from the first and second stages:

$$e_o^{1,2*} = e_o^{1*} e_o^{2*} \tag{4}$$

where $e_o^{1,2*}$ is the overall efficiencies of DMU_o.

4.3 ANNs

To examine the impacts of exogenous variables, it has been a common practice to regress DEA estimates against the variables. However, how to choose the most appropriate regression models – since DEA scores can fall between zero and unity – is still an ongoing discussion (see, e.g. McDonald, 2009; Ramalho *et al.*, 2010; Simar and Wilson, 2011).

As an alternative, ANN-based solutions have been proposed. ANNs are powerful tools for solving complex problems in several fields of application, such as classification, function approximation, optimization, data processing, and control systems (Dias and Silvestre, 2011; Mohaghegh *et al.*, 1995). It is beyond the scope of the present paper to give an in-depth description of this method; interested readers are instead referred to standard introductory textbooks on neural networks (e.g. Rojas, 1996).

The application of ANNs is not without its drawbacks. ANNs behave, essentially, as black boxes, providing no insight into the roles or relative importance of factors (Lee *et al.*, 2004); furthermore, the reasoning process is not transparent to users (Marsh and Fayek, 2010).

Neural networks do, however, have coefficients, which are the interconnection weights between input and output layers (Ratner, 2003). There have been formal efforts to translate connection weights into coefficient-like information, such as the Garson (1991) algorithm. Olden and Jackson (2002), however, demonstrated that this algorithm can be potentially misleading when interpreting input variable contributions, and they therefore propose instead using their connection-weight approach, which has outperformed all other approaches with regard to quantifying variable importance (Olden *et al.*, 2004).

Unlike Garson, who used absolute connection weights, Olden and Jackson used the original weights. Mathematically, the importance of an input variable under the connection-weight approach can be written as follows (Kemp *et al.*, 2007):

$$I(i) = \sum_{x=1}^{k} CW_{ih(x)} CW_{ho(x)} \tag{5}$$

where $I(i)$ is the relative importance of input i, k, the total number of hidden nodes; x, the index number of hidden node, $CW_{ih(x)}$, the connection weight between input i and hidden node x, and $CW_{ho(\mathrm{x})}$, the connection weight between hidden node x and output node. Equation (5) can be normalized into:

$$I_N(i) = \frac{|I(i)|}{\sum_{i=1}^{p} |I(i)|} \tag{6}$$

where $I_N(i)$ is the normalized relative importance of input i, and p, the total number of input nodes. This approach relies upon the fact that the connection weights between neurons are the linkages between the inputs and the output of the network, and therefore also the links between the problem and the solution. As a result, the relative contribution of each independent variable to the predictive output of the neural network depends primarily on the magnitude and direction of these connection weights (Olden, 2010).

4.4 Input-output measures and exogenous variables

Our model consisted of one set of native input measures and two sets of output measures: technical and financial. Input and output measures, as well as exogenous variables, were selected both based on their relevance to water operation and on data availability. Choosing input-output variables is a matter of researcher judgment, so long as inputs sufficiently reflect the resource used and outputs sufficiently reflect the volume and quality of services encapsulated in the function being modeled (Thanassoulis, 2000a). It is thus possible that the selected indicators may differ from those commonly used. Data availability were included as a criterion because it allowed us to add many MWUs of interest. Collecting vast numbers of data sets from hundreds of units would be difficult and costly, both in terms of time and money.

We identified the following native input measures: the number of staff per 1,000 connections, the OM costs per 1,000 connections, and the ratio of overhead (OH) costs to revenues. The number of connections could be seen as a desirable output – indeed, Berg and Lin (2008) used it as input in their SFA and DEA models but it has been common practice to instead use this number as a divisor. This ensures a reasonable comparison between utilities with different sizes of operation and staff productivity.

As MWUs are government owned, overstaffing may be a serious problem due to weak managerial incentives for cost containment and interference from local politicians (Lin, 2005). Following Romano and Guerrini (2011), the present study also considered the cost data as inputs. While the data for number of staff members, number of connections, and OM costs were available, those for OH costs and revenue were not (although the ratios were known).

We chose accounted-for water level, ratio of service coverage, ratio of volume of water produced to volume of water distributed, operating hours per day, and billing effectiveness rate as the intermediate outputs. Accounted-for water level is the complement of non-accounted-for water level, and was simply calculated by subtracting the latter from unity. Operating hours per day were obtained by dividing the total operating hours for one year by 365, and billing effectiveness rate was computed by dividing the number of bills paid by customers by the total number of bills. These selected outputs were the same as those used in Braadbaart (2007), Lin (2005), and Romano and Guerrini (2011).

To include financial performance in the model, we set return on equity (ROE), ratio of operating revenue to operating costs, cash ratio, and solvability as output measures. While ratios of operating revenue to operating costs were available, operating revenue and operating cost data sets were not in the database. Had complete data been available for revenues and costs, we could certainly have reduced the number of inputs and outputs in our DEA system.

An MWU usually manages several units of water services, which may cover both urban and rural areas. We introduced the urban ratio into the system to see if the geographical areas being served had any impact on efficiency. This ratio was expressed as the number of units serving urban areas divided by total number of units.

The ratio of average tariff to base production costs was assumed to be exogenous because tariffs are set externally by local governments, upon the approval of local parliaments. The size of population served was classified as an exogenous variable as it is likely to be outside of managerial control (Thanassoulis, 2000a). We also added the ratio of the number of trained staff to the total number of staff. Despite the fact that staff makeup was under the control of MWU management, this ratio was regarded as neither an input nor an output; we therefore instead explored whether or not it had any effect on efficiency.

Our exogenous environmental variables thus comprised the ratio of average tariffs (per m^3) to base production costs (per m^3), size of population served, domestic water consumption per month per connection, urban ratio, and per-capita municipal gross domestic product (GDP). Domestic water consumption was calculated as the average volume of water sold for domestic use per month, divided by the total number of domestic connections.

4.5 Data collection

Our required data were mainly collected from the Supporting Agency for the Development of Drinking Water Provision System (henceforth "The Supporting Agency"). This agency has routinely reported on the annual performance of Indonesia's MWUs; this information is made public for accountability purposes under the Supporting Agency's web site (www.bppspam.com). The latest report, published in 2012, covered the years of 2006-2010, and included 335 operators throughout the country (2012).

We also benefited from the rich web database SIMSPAM (Information Management System for Drinking Water Development). Administered by the Directorate of Drinking Water Development of the Ministry of Public Works, it encompasses data from all of Indonesia's MWUs. The data from the SIMSPAM system were used to supplement the Supporting Agency's data.

It should be noted that the Supporting Agency calculated solvability as the ratio of total assets to total liabilities, which is the reciprocal of the standard solvency ratio. Also, based on assessment guidance provided by the Finance and Development Supervisory Agency (FDSA, 2011), any negative ROE was set as zero, which might not reflect the real situation. However, this method was appropriate for the present study given that the standard leader-follower DEA model does not allow for negative output measures.

Because it was not the objective of this study to evaluate changes in the efficiency of MWUs over time (window analysis), we used only the latest figures (2010 data) for analysis. The 2013 edition was recently released; however, it is not as detailed as the 2010 version, and is more of a summary of performance attained. Prepared based on the audits of *Badan Pengawasan Keuangan dan Pembangunan* (Finance and Development Supervisory Agency), those reports are the most complete and reliable data sources regarding the performance of Indonesian MWUs to date. The individual data regarding 2010 municipal per-capita GDP were collected from Statistics Indonesia (www.bps.go.id).

One serious problem we encountered during data collection from both of our data sources were a substantial number of missing or unreasonable data items across multiple variables. Although it would be computationally possible to include these data, the DEA would produce unreliable scores, as empty data would be treated as zeros. To deal with this, we omitted samples that had data missing for one or more variables, at the cost of a considerable reduction in our sample size. Of 335 (or 483, if using the SIMSPAM system), only 269 samples were valid for evaluation after applying this criterion. However, despite this reduction in sample size, the sample still well exceeds the guidelines regarding minimum allowable sample size (twice the product of the number of inputs and number of outputs, or three times the sum of the number of inputs and number of outputs; Avkiran, 2011). Table I shows the descriptive statistics of input, intermediate output, and output data.

4.6 Data analysis

We evaluated the stage and overall efficiencies of the two models. The first model viewed serviceability as the leader and profitability as the follower, and the second model assumed the converse: profitability as the leader and serviceability as the follower. Due to the limited length of this paper, we will put greater emphasis on the first model in the discussion, as we deemed serviceability to be of greater public interest. In the first model we first used Equation (2) to optimize the efficiency score of technical outputs to inputs (the leader) for the DMU of interest. Next, we used Equation (3) to maximize the efficiency score of financial outputs to technical outputs (the follower) on the condition that the leader's score remained unchanged. The computational procedures for the second model are identical to those used for the first model.

To solve linear programming (LP) problems, we used OpenSolver, a free Excel® add-on software package (opensolver.org). To deal with the cumbersome repetitive processes of computation required with manual execution, we coded a simple Visual

Measure	Minimum	Maximum	Mean	SD
	Statistics			
Input				
Number of staff per 1,000 connections	0.28	41.97	8.74	5.82
Ratio of OM cost to 1,000 connections (IDR million)	0.72	26,677.16	700.21	1,770.21
Ratio of OH cost to revenue	0.08	1.61	0.58	0.25
(Intermediate) output				
Accounted for water	0.21	0.90	0.67	0.13
Coverage areas of service	0.03	0.99	0.37	0.21
Production efficiency	0.07	1.57	0.87	0.18
Operating hours	3.00	24.00	18.71	5.69
Billing effectiveness	0.39	1.00	0.88	0.15
Return on equity	0.00	0.73	0.05	0.10
Ratio of revenue to operating cost	0.10	1.67	0.87	0.26
Cash ratio	0.00	1,036.17	13.31	73.07
Solvability	0.00	3,387.05	36.20	219.13
Exogenous factor				
Tariff ratio	0.25	8.19	1.70	1.06
Population size	1,070.00	7,342,420.00	123,992.53	468,877.35
Domestic water consumption	5.21	33.98	17.28	4.39
Urban ratio	0.00	1.00	0.64	0.40
Per-capita GDP (IDR million)	1,200.00	369,510.00	17,983.99	28,188.85
Training ratio	0.00	1.99	0.13	0.23

Table I. Descriptive statistics of input, (Intermediate) output, and exogenous factors

Basic for Applications script, with reference to OpenSolver, to automate the calculations. Once the optimal solutions were obtained, the overall efficiencies were then computed using Equation (4). Data analysis for ANNs were supported using Statistical Product and Service Solutions Statistics® R. 21. In total, 70 percent of the total samples were designated as the training set, and the remaining 30 percent was designated as a testing set.

5. Results and discussion

When serviceability was treated as the leader (Model 1), the LP solutions yielded the first-stage DEA scores. These scores ranged from 16 to 100 percent, with a mean of 55 percent. The scores for the second stage ranged from 0 to 100 percent, with a mean of 32 percent. The overall efficiencies were between 0 and 100 percent, with a mean of 20 percent. These levels of performance indicate sizable opportunities for improvement. A total of 19 and five DMUs had first- and second-stage DEA scores of unity, respectively, but there was only one performer with a perfect overall score of 1.00 (MWU Banyuwangi, see Table II). Again, due to the length constraint on the present paper, Table II only presents the 20 best performers based on overall efficiency scores.

The sample distribution of the first- and second-stage DEA scores was asymmetrical, and tended to skew to the right, indicating that more operators exhibited below-average performance (see Figures 2 and 3). The calculated skewness coefficients of the first stage, the second stage, and the overall scores were 0.68, 0.72, and 1.73, respectively. The Kolmogorov-Smirnov test further confirmed that the null hypothesis of normal distribution for the scores was rejected at the 0.05 level for all efficiency measures. Such poor average scores should attract considerable

Rank	Municipal water utilities	Model 1 First stage	second stage	Overall	Municipal water utilities	Model 2 First stage	Second stage	Overall
1	Banyuwangi	1.00	1.00	1.00	Balikpapan	1.00	0.50	0.50
2	Mataram	1.00	0.98	0.98	Jakarta	1.00	0.47	0.47
3	Sragen	1.00	0.87	0.87	Yogyakarta	1.00	0.43	0.43
4	Lubuk Linggau	0.86	1.00	0.86	Bogor	1.00	0.37	0.37
5	Banjar	1.00	0.86	0.86	Kuningan	1.00	0.28	0.28
6	Yogyakarta	1.00	0.86	0.86	Malang	1.00	0.26	0.26
7	Balikpapan	1.00	0.85	0.85	Buleleng	1.00	0.26	0.26
8	Jombang	0.82	1.00	0.82	Banjar	1.00	0.25	0.25
9	Indramayu	1.00	0.80	0.80	Sragen	1.00	0.23	0.23
10	Tuban	1.00	0.80	0.80	Mataram	1.00	0.21	0.21
11	Bogor	1.00	0.79	0.79	Lubuk Linggau	1.00	0.18	0.18
12	Jepara	1.00	0.75	0.75	Magelang City	0.70	0.22	0.15
13	Malang	1.00	0.74	0.74	Tuban	0.84	0.18	0.15
14	Jakarta	1.00	0.74	0.74	Gunung Kidul	0.71	0.21	0.15
15	Buleleng	1.00	0.73	0.73	Salatiga	0.67	0.22	0.15
16	Kuningan	1.00	0.69	0.69	Banjarmasin	0.50	0.29	0.14
17	West Halmahera	0.83	0.79	0.66	Pekalongan	0.73	0.19	0.14
18	Surabaya	0.92	0.71	0.65	Surabaya	0.99	0.14	0.14
19	Klungkung	0.87	0.75	0.65	Sleman	0.59	0.23	0.13
20	Gunung Kidul	1.00	0.60	0.60	Banyuwangi	1.00	0.13	0.13

Table II. Top 20 most efficient municipal water utilities

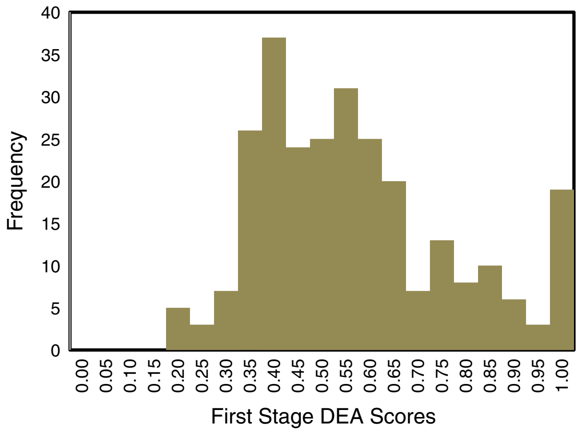

Figure 2. Sample distribution of the first stage DEA scores

attention from the GoI, as there are a large number of operators that are having serious efficiency problems.

Further pooled data indicated efficiency discrepancies among regions. The Supporting Agency clusters the regions into four large regions covering Indonesia's main islands: Region I (Sumatera), Region II (Java), Region III (Kalimantan and Sulawesi), and Region IV (Bali, West Nusa Tenggara, East Nusa Tenggara, Maluku, and Papua). The average overall DEA scores for these regions are as follows: Region I = 11 percent, Region II = 28 percent, Region III = 16 percent, Region IV = 27 percent. It is no surprise that

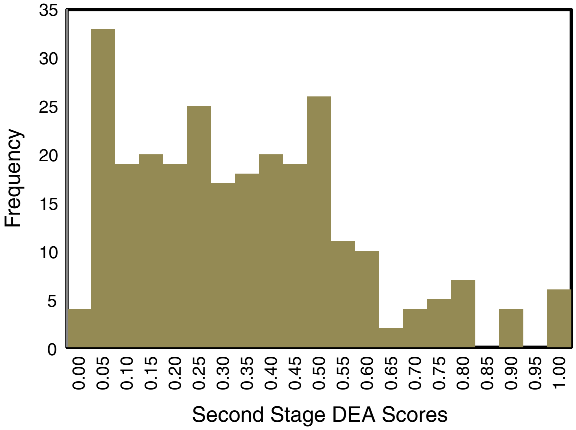

Figure 3. Sample distribution of the second stage DEA scores

Region II, which contains Java Island, the country's most economically developed and populous island, outperforms all other regions. However, when we explored by individual island, Bali Island, one of the most famous world tourism destinations, ranks first, with the average DEA scores in the first stage (77 percent), the second stage (54 percent), and overall stages (43 percent), which are well above the national average.

Figure 4 displays the scatter plot of the first- and second-stage scores using two-dimensional Cartesian coordinates for Model 1, with the former denoted as the axis and the latter as the ordinate, to show the relationship between the two types of scores. The crosshairs, placed on the corresponding mean scores, divide the plot into four quadrants. Starting from the upper right corner and moving counterclockwise, these quadrants are Quadrant I, indicating high first-stage scores and high second-stage scores, Quadrant II (low first-stage scores and high second-stage scores), Quadrant III

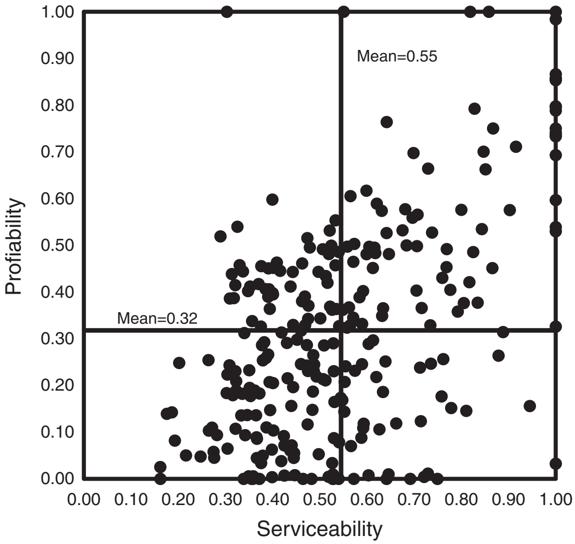

Figure 4. Scatter diagram of efficiencies in serviceability and profitability (Model 1)

(low first-stage scores and low second-stage scores), and Quadrant IV (high first-stage scores and low second-stage scores).

Table III presents the subset means of DEA for the two stages, inputs, intermediate outputs, and outputs based on this quadrant analysis of Model 1. As expected, the calculated DEA scores are highly consistent with pooled performance data; i.e., lower inputs lead to greater outputs: MWUs that fell into Quadrant I outperformed those MWUS in other quadrants for most variables, and MWUs in Quadrant III tended to underperform relative to their peers. The performance of MWUs in Quadrants II and IV was relatively comparable, and fell between the performances of MWUs in Quadrants III and I.

Placement in these different quadrants implies different appropriate management actions. For instance, MWUs in Quadrant IV need to improve their performance in second-stage efficiencies while maintaining their already above-par performance in first-stage efficiencies. Looking at Figure 4, urgent actions to improve performance for both stages are imperative for the 104 MWUs (39 percent of the total sample) that fell into Quadrant III. This subset had average first- and second-stage scores of only 39 percent (minimum = 16 percent, maximum = 54 percent) and 14 percent (minimum = 0 percent, maximum = 32 percent), respectively.

5.1 Model 2: profitability as the leader

When we developed another situation, with profitability as the leader, we found a first stage average efficiency of 40 percent, second stage efficiency of 22 percent, and overall efficiency as low as 7 percent. The skewness of data also exhibits a substantial increase if compared to that of Model 1: 1.07 for the first stage, 2.83 for the second stage, and 3.76 for overall stage. Table II displays the top 20 performers based on their overall

	Quadrant			
Variables	I ($n = 76$)	II ($n = 52$)	III ($n = 104$)	IV ($n = 37$)
DEA scores				
First stage	0.77	0.43	0.39	0.68
Second stage	0.58	0.44	0.14	0.14
Overall	0.46	0.19	0.05	0.10
Inputs				
Number of staff per 1,000 connections	5.22	9.92	11.47	6.67
Ratio OM cost to 1,000 connections (IDR million)	416.33	1,491.80	633.08	359.52
Ratio of OH cost to revenue	0.39	0.57	0.77	0.43
Intermediate outputs				
Accounted for water	0.73	0.69	0.62	0.65
Coverage areas of service	0.44	0.36	0.30	0.40
Production efficiency	0.90	0.89	0.84	0.85
Operating hours	20.79	17.19	17.66	19.54
Billing effectiveness	0.93	0.84	0.85	0.93
Outputs				
Return on equity	0.10	0.04	0.01	0.05
Ratio of revenue to operating cost	1.07	0.90	0.67	0.95
Cash ratio	24.43	15.52	7.26	4.35
Solvability	38.03	92.37	16.22	9.69

Table III. Means of DEA scores, input, intermediate, and output measures by quadrant (Model 1) efficiencies, if profitability is set as the leader. These results went in the same direction of that of Model 1, as indicated by a strong positive correlation between the efficiencies of Models 1 and 2 ($r = 0.701$, significant at the 0.001 level). This finding implies that an MWU that operates efficiently according to Model 1 will tend to operate efficiently in Model 2, although at a different magnitude, and vice versa. When using Model 2, the number of MWUs located in Quadrant I declines markedly, from 76 to only 17, shifting to other quadrants. We may therefore conclude that prioritizing profitability as an output would result in lower overall efficiency, indicating more formidable challenges ahead – this would have major implications for GoI policy. Owing to the fact that MWUs should also carry some social responsibility, adopting the mission of ensuring universal water access, Model 2 was deemed difficult to apply in most cases. However, together these findings offer different perspectives on MWU efficiencies; these different findings could be used to establish the need for improvements if profitability leads to serviceability. This diversity of perspectives is the distinctive advantage of using the Stackelberg DEA model.

5.2 Importance of exogenous variables

The multilayer perceptron method and the back-propagation algorithm for data training were used to obtain the connection weights. To ensure that the predicted values of ANNs would remain within the interval of 0-1, a sigmoid activation function was adopted. Table IV presents an example of ANN synaptic weights that connect the nodes of exogenous variables through one hidden layer to the node of the first-stage DEA scores of Model 1. Using Olden's (2010) recommendations to interpret Table IV, input variables with larger connection weights represent greater intensities of signal transfer and therefore are more important in predicting the output than are variables of smaller weights. Negative signs represent inhibitory effects on neurons (i.e. reducing the intensity or contribution of the incoming signal and thus negatively affecting the output, e.g., the connection weight of urban ratio to the hidden layer) and positive signs represent excitatory effects on neurons (i.e. increasing the intensity of the incoming signal and positively affecting the output; e.g., the connection weight of domestic water consumption to the hidden layer; Olden, 2010).

Table IV. Example of synaptic weights: Stage 1 efficiency (Model 1)

	Predicted				
	Hidden layer 1				
Exogenous variable	H(1:1)	H(1:2)	H(1:3)	H(1:4)	Output layer
Bias	0.421	0.362	−0.420	−0.235	
Urban ratio	−0.252	0.550	0.224	−0.181	
Per-capita GDP (IDR million)	−0.339	0.084	−0.226	−0.123	
Tariff ratio	−0.393	−0.181	0.205	−0.192	
Population size	0.398	−0.644	0.190	−0.947	
Domestic water consumption	0.443	−0.332	−0.327	0.131	
Training ratio	−0.302	−0.523	−0.155	−0.020	
(Bias)					−0.115
H(1:1)					0.324
H(1:2)					−0.401
H(1:3)					0.175
H(1:4)					−0.921

Using Equations (5) and (6), the individual relative importance of exogenous variables for each stage's efficiency can be estimated for both models (see Table V). One conclusion that can be made based on the data presented in this table is that the size of population served is the most important exogenous environmental variable in explaining MWU efficiencies for both Models 1 and 2.

We initially had expected that the urban ratio, as well as other variables, would be significantly correlated with efficiencies. However, our analysis did not affirm this expectation. For instance, we assumed that urban MWUs operate more efficiently than rural ones, but, in fact, no substantial difference in efficiencies was observed between the two groups. The high importance of the size of population served led us to conduct a more detailed analysis of how scores change with the urban ratio, using a simple correlation analysis between efficiencies and the sizes of population.

As expected, a consistent increase in DEA scores accompanied increases in the size of population served, with coefficients of correlation of 0.530 for Model 1 and 0.471 for Model 2 (both significant at least at the 0.001 level). One plausible explanation for this effect may be economy of scale. Economies of scale refer to cases for which all inputs – not just machinery and factory space, but also labor and other inputs – are adjusted in fixed proportions. If quantities of some inputs are increased by relatively greater amounts than others, then a size rather than scale relationship is observed (Dayananda *et al.*, 2002). The findings of the present study would then affirm those of da Motta and Moreira (2004) and Romano and Guerrini (2011).

Zschille and Walter (2010) provided a brief overview of international studies on economies of scale and density, with a focus on Europe. These studies indicated that there are at least two economies of scale for water systems: for capital equipment – probably the most familiar to readers – as well as for many ordinary business operations, such as

	Efficiencies					
	Stage 1		Stage 2		Overall	
Exogenous variable	Connection weight	Normalized importance (%)	Connection weight	Normalized importance (%)	Connection weight	Normalized importance (%)
Model 1						
Urban ratio	−0.096	5.29	−0.117	8.58	−0.219	14.26
Per-capita GDP (IDR million)	−0.070	3.83	−0.064	4.65	0.022	1.46
Tariff ratio	0.158	8.70	−0.018	1.32	−0.018	1.17
Population size	1.292	71.07	0.829	60.64	0.948	61.73
Domestic water consumption	0.099	5.43	0.308	22.55	0.251	16.33
Training ratio	0.103	5.67	0.031	2.26	0.077	5.05
Model 2						
Urban ratio	−0.096	4.70	0.043	4.67	−0.113	8.40
Per-capita GDP (IDR million)	−0.327	15.98	0.058	6.27	−0.122	9.04
Tariff ratio	0.096	4.69	−0.452	49.22	−0.205	15.20
Population size	1.179	57.55	−0.261	28.49	0.630	46.79
Domestic water consumption	0.263	12.83	−0.051	5.56	0.211	15.67
Training ratio	0.087	4.24	−0.053	5.78	0.066	4.90

Table V. ANN-based exogenous variable importance billing, purchasing, and computer systems. This second category also includes ancillary water treatment and testing operations, two factors relevant to the context of the present paper (Shih *et al.*, 2004). Capital investment theory offers a six-tenth-factor rule of thumb, equating the ratio of the total cost of two similar plants of different capacities to the ratio of the capacities raised to a power, which typically ranges between 0.6 and 0.9 (Lang and Merino, 1993). If this rule also applies to water treatment plants, then an increased capacity would result in a reduction in the cost of every unit.

5.3 Remedial actions

A wide range of efforts can be initiated to help improve the efficiencies of these MWUs, from strengthening their technical and financial capacities to introducing PPPs under a B-B model. Skills and professional enhancement can be pursued through capacity building programs, which have been routinely managed by central and local governments. However, these programs need to be based on detailed need assessments, and effectiveness would need to be regularly monitored and evaluated. Capacity improvement cannot be changed overnight, and ongoing and continuous programs, as opposed to project-based or one-off measures, are essential. The GoI may also facilitate on-the-job trainings to allow knowledge transfer between key staff in highly performing and poorly performing MWUs of comparable sizes. Open recruitment for top-level functions could also be an effective way to overcome insufficient professional competence.

The GoI has also embarked upon a debt write-off and restructuring plan to help alleviate the chronic financial troubles of many distressed MWUs; however, not all of these MWUs were willing to participate in the plan, as it entailed a set of commitments they had to meet. From a financial standpoint, the next and even more urgent issue is the excessively low water charges set by local authorities; these charges are insufficient for OM cost, let alone opportunity cost of capital. While tariff setting remains the responsibility of local governments, governments are often reluctant to incur popular resentment due to price increases, even if these increases are necessary to cover costs (Eberhard, 2007). Cost recovery policies for public utilities may therefore be a simple and effective solution for overly low water charges, although such policies must be accompanied by good corporate governance of water services.

The PPP model could be a viable means to help the public sector benefit from the private sector's competitive advantages, improving the water supply sector and helping to push for water policies that benefit the nation's poor (Ameyaw and Chan, 2013; Bakker, 2007; Choi *et al.*, 2010; Ke *et al.*, 2010). However, the application of this approach is not without challenges. There is fierce public opposition to water commercialization, as well as widespread fear that private companies would abuse their monopoly powers; these fears would only increase if customers are accustomed to paying unrealistically low rates. To make PPPs work in the water sector, certain critical success factors must be present, including project profitability, asset quality, fair risk allocation, competitive tendering, internal coordination within the government, employment of professional advisors, corporate governance, and government supervision (Meng *et al.*, 2011).

The presence of economies of scale in water treatment plants would justify larger, more integrated networks of smaller water utilities. The idea of restructuring the water sector through the integration of smaller networks to improve efficiencies deserves serious consideration. Under the existing structure, the areas serviced by individual public water utilities are demarcated by administrative borders. It would be interesting to explore the benefit and cost if the establishment of MWUs was more watershed- than district-oriented. Municipal governments sharing the same watershed might find it helpful to enter into municipal – municipal government cooperative agreements to improve efficiency. However, further investigation is needed to confirm this argument, as well as to explore how such an integration might be achieved.

6. Concluding remarks

We have presented a yardstick comparison of efficiencies for 269 of Indonesia's MWUs, under the framework of a two-stage Stackelberg leader-follower DEA model. A substantial number of MWUs were found to be operating at inefficient levels as compared to their peers. Our analysis indicates that systematic and comprehensive measures to enhance the country's water-sector efficiencies are urgently needed, especially if the GoI is concerned about achieving MDGs. Due to the limited resources of both central and municipal governments, the highest priority for technical and financial assistance should be assigned to MWUs suffering from below-average efficiencies for both serviceability and profitability. The next-highest priority should be given to MWUs that are performing poorly at either for either serviceability or profitability.

To date, Indonesia has not implemented a performance-indexed incentive mechanism in the water sector, through either yardstick or benchmarking regulations. If peer comparison is deemed too threatening, then performance can at least be benchmarked to previous assessments of a given MWU, necessitating continuous improvement. However, we recommend conducting additional research to determine the best method by which Indonesia can push for greater efficiency in its water sector.

The DEA scores resulting from the present study could also be fed into a discriminant model to distinguish between MWUs that pose high and low credit risks. At present, a total of 110 out of 178 MWUs that borrowed from the GoI defaulted on their payments, and an improved model could certainly help the central government to evaluate loans proposed by individual MWUs. The central government could also use this information to quantify the contingent liabilities associated with providing debt guarantees to MWUs seeking loans from commercial banks for business expansion. Such guarantees have been offered by the GoI since 2009, with the goal of accelerating public water provision, and the inclusion of DEA scores as one of the predictors could improve the GoI's accuracy in quantifying their liability.

Our study was also limited by its assumption that all MWUs operate with CRS. Although overall acceptable (see Thanassoulis, 2000b for a more in-depth discussion of the debate as to whether a CRS or VRS assumption is appropriate in regulatory contexts such as the water sector), this assumption might be regarded as overly restrictive (Ramanathan, 2003). Relaxing the assumption to VRS would likely improve the scores of inefficient MWUs (Chen *et al.*, 2009), especially when regulators may indirectly control the scale size of MWUs by permitting mergers and acquisitions of regulated firms, as proposed in this paper, given that the scale size is dependent on population served and not controllable by utilities. It would therefore be advisable to scrutinize the effect of introducing an assumption of VRS into this model. Liang *et al.* (2008) model, used in this study, that employs a multiplicative-based approach is not readily applicable for VRS settings. In that VRS situation, one should instead modify the original model or, alternatively, return to additive-based models (e.g. Chen *et al.*, 2009 and Chen and Zhu, 2004). This paper is no way intended to discuss how MWU efficiencies compare under the assumptions of CRS or VRS; rather, this issue is to be left as an open question for future research.

References

Ameyaw, E.E. and Chan, A.P.C. (2013), "Identifying public–private partnership (PPP) risks in managing water supply projects in Ghana", *Journal of Facility Management*, Vol. 11 No. 2, pp. 152-182.

Avkiran, N.K. (2011), "Applications of data envelopment analysis in the service sector", in Cooper, W.W., Seiford, L.M. and Zhu, J. (Eds), *Handbook on Data Envelopment Analysis*, Springer, New York, NY, pp. 403-444.

Bakker, K.J. (2007), "Trickle down? Private sector participation and the pro-poor water supply debate in Jakarta, Indonesia", *Geoforum*, Vol. 38 No. 5, pp. 855-868.

Barber, E. (2004), "Benchmarking the management of projects: a review of current thinking", *International Journal of Project Management*, Vol. 22 No. 4, pp. 301-307.

Bell, R.A. and Morey, R.C. (1994), "The search for appropriate benchmarking partners: a macro approach and application to corporate travel management", *Omega International Journal of Management Science*, Vol. 22 No. 5, pp. 477-490.

Berg, S. and Lin, C. (2008), "Consistency in performance rankings: the Peru water sector", *Applied Economics*, Vol. 40 No. 6, pp. 793-805.

Braadbaart, O. (2007), "Collaborative benchmarking, transparency and performance: evidence from the Netherlands water supply industry", *Benchmarking: An International Journal*, Vol. 14 No. 6, pp. 677-692.

Camp, R.C. (1989), *Benchmarking–The Search for Industry's Best Practices that Lead to Superior Performance*, Quarterly Press, Milwaukee, WI.

Camp, R.C. (1992), "Learning from the best leads to superior performance", *Journal of Business Strategy*, Vol. 13 No. 3, pp. 3-6.

Chen, Y. and Zhu, J. (2004), "Measuring information technology's indirect impact on firm performance", *Information Technology and Management*, Vol. 5 Nos 1-2, pp. 9-22.

Chen, Y., Cook, W.D., Li, N. and Zhu, J. (2009), "Additive efficiency decomposition in two-stage DEA", *European Journal of Operational Research*, Vol. 196 No. 3, pp. 1170-1176.

Choi, J.-H., Chung, J. and Lee, D.-J. (2010), "Risk perception analysis: participation in China's water PPP market", *International Journal of Project Management*, Vol. 28 No. 6, pp. 580-592.

Collier, T., Johnson, A.L. and Ruggiero, J. (2011), "Technical efficiency estimation with multiple inputs and multiple outputs using regression analysis", *European Journal of Operational Research*, Vol. 208 No. 2, pp. 153-160.

Cook, W.D., Liang, L. and Zhu, J. (2011), "Modeling DMU's internal structures: cooperative and noncooperative approaches", in Cooper, W.W., Seiford, L.M. and Zhu, J. (Eds), *Handbook on Data Envelopment Analysis*, Springer, New York, NY, pp. 297-314.

Corton, M.L. (2003), "Benchmarking in the Latin American water sector: the case of Peru", *Utilities Policy*, Vol. 11 No. 3, pp. 133-142.

Corton, M.L. and Berg, S.V. (2009), "Benchmarking central American water utilities", *Utilities Policy*, Vol. 17 No. 3, pp. 267-275.

da Motta, R.S. and Moreira, A. (2004), "Efficiency and regulation in the sanitation sector in Brazil", *Utilities Policy*, Vol. 14 No. 3, pp. 185-195.

Dayananda, D., Irons, R., Harrison, S., Herbohn, J. and Rowland, P. (2002), *Capital Budgeting: Financial Appraisal of Investment Projects*, Cambridge University Press, Cambridge.

Dias, J.L.R. and Silvestre, N. (2011), "A neural network based closed-form solution for the distortional buckling of elliptical tubes", *Engineering Structures*, Vol. 33 No. 6, pp. 2015-2024.

Eberhard, A. (2007), "Matching regulatory design to country circumstances: the potential of hybrid and transitional models", *Gridlines (Note 23), Public-Private Infrastructure Advisory Facility*, World Bank, Washington, DC, available at: www.gsb.uct.ac.za/files/Grid-23-Hybrid-31.pdf (accessed September 29, 2014).

Finance and Development Supervisory Agency (2011), *Auditing Guidance for Municipal Water Utility Performance*, FDSA, Jakarta.

Garson, G.D. (1991), "Interpreting neural-network connection weights", *AI Expert*, Vol. 6 No. 4, pp. 46-51.

Hassanein, A.A.G. and Khalifa, R.A. (2007), "Financial and operational performance indicators applied to public and private water and wastewater utilities", *Engineering, Construction and Architectural Management*, Vol. 14 No. 5, pp. 479-492.

Hon, L.Y. and Lee, C. (2009), "Efficiency in the Malaysian water industry: A DEA and regression analysis", available at: https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name = SERC2007&paper_id = 109 (accessed November 18, 2013).

Janakanarajan, S., Llorente, M. and Zérah, M.-H. (2006), "Urban water conflicts in Indian cities: man-made scarcity as a critical factor", available at: http://r4d.dfid.gov.uk/PDF/Outputs/Water/R8324-UNESCOpap.pdf (accessed December 18, 2013).

Ke, Y., Wang, S.Q., Chan, A.P.C. and Lam, P.T.I. (2010), "Preferred risk allocation in China's public-private partnership (PPP) projects", *International Journal of Project Management*, Vol. 28 No. 5, pp. 482-492.

Kemp, S.J., Zaradic, P. and Hansen, F. (2007), "An approach for determining relative input parameter importance and significance in artificial neural networks", *Biological Modelling*, Vol. 204 Nos 3/4, pp. 326-334.

Lang, H.J. and Merino, N. (1993), *The Selection Process for Capital Projects*, John Wiley & Sons, New York, NY.

Lee, M.-J., Hanna, A.S. and Loh, W.-Y. (2004), "Decision tree approach to classify and quantify cumulative impact of change orders on productivity", *Journal of Computing in Civil Engineering*, Vol. 18 No. 2, pp. 132-144.

Liang, L., Cook, W.D. and Zhu, J. (2008), "DEA models for two-stage processes: game approach and efficiency decomposition", *Naval Resources Logistic Quarterly*, Vol. 55 No. 7, pp. 643-653.

Liang, L., Yang, F., Cook, W.D. and Zhu, J. (2006), "DEA models for supply chain efficiency evaluation", *Annals of Operations Research*, Vol. 145 No. 1, pp. 35-49.

Lin, C. (2005), "Service quality and prospects for benchmarking: evidence from the Peru water sector", *Utilities Policy*, Vol. 13 No. 2, pp. 230-239.

Love, R., Bunney, H.S., Smith, M. and Dale, B.G. (1998), "Benchmarking in water supply services: the lessons learnt", *Benchmarking for Quality Management & Technology*, Vol. 5 No. 1, pp. 59-70.

Luu, V.T., Kim, S.-Y. and Huynh, T.-A. (2008), "Improving project management performance of large contractors using benchmarking approach", *International Journal of Project Management*, Vol. 26 No. 7, pp. 758-769.

McDonald, J. (2009), "Using least squares and tobit in second stage DEA efficiency analyses", *European Journal of Operational Research*, Vol. 197 No. 2, pp. 792-798.

Macpershon, A.J., Principe, P.P. and Shao, Y. (2013), "Controlling for exogenous environmental variables when using data envelopment analysis for regional enviromental assessments", *Journal of Environmental Management*, Vol. 119, April, pp. 220-229.

Marques, R., Berg, S. and Yane, S. (2014), "Nonparametric benchmarking of Japanese water utilities: institutional and enviromental factors affecting efficiency", *Journal of Water Resource Planning and Management*, Vol. 140 No. 5, pp. 562-571.

Marsh, K. and Fayek, A.R. (2010), "Suretyassist: fuzzy expert system to assist surety underwriters in evaluating construction contractors for bonding", *Journal of Construction Engineering and Management*, Vol. 136 No. 11, pp. 1219-1226.

Meng, X., Zhao, Q. and Shen, Q. (2011), "Critical success factors for transfer-operate-transfer urban water supply projects in China", *Journal of Construction Engineering and Management*, Vol. 27 No. 4, pp. 243-251.

Ministry of Public Works (2010), *Strategic Plan 2010-2014*, Ministry of Public Works, Jakarta.

Mohaghegh, S., Arefi, R., Ameri, S. and Rose, D. (1995), "Design and development of an artificial neural network for estimation of formation permeability", *SPE Computer Applications*, Vol. 7 No. 6, pp. 151-154.

Mugisha, S. (2014), "Frontier distance function analysis for water supply systems", *Municipal Engineer*, Vol. 167 No. 1, pp. 11-21.

Norton, J.W. and Weber, W.J. (2009), "Water utility efficiency assessment using a data envelopment analysis procedure", *Journal of Infrastructure Systems*, Vol. 15 No. 2, pp. 80-87.

Olden, J.D. (2010), "Current velocity shapes the functional connectivity of benthiscapes to stream insect movement", in Tosh, C.R. and Ruxton, G.D. (Eds), *Modelling Perception with Artificial Neural Networks*, Cambridge University Press, Cambridge, pp. 351-373.

Olden, J.D. and Jackson, D.A. (2002), "Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks", *Ecological Modelling*, Vol. 154 Nos 1/2, pp. 135-150.

Olden, J.D., Joy, M.K. and Death, R.G. (2004), "An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data", *Ecological Modelling*, Vol. 178 Nos 3/4, pp. 389-397.

Ramalho, E.A., Ramalho, J.J.S. and Henriques, P.D. (2010), "Fractional regression models for second stage DEA efficiency analyses", *Journal of Productivity Analysis*, Vol. 34 No. 3, pp. 239-255.

Ramanathan, R. (2003), *An introduction to Data Envelopment Analysis: A Tool for Performance Measurement*, Sage Publications, New Delhi.

Ratner, B. (2003), *Statistical Modeling and Analysis for Database Marketing: Effective Techniques for Mining Big Data*, Chapman & Hall/CRC, New York, NY.

Rojas, R. (1996), *Neural Networks: A Systematic Introduction*, Springer, Heidelberg.

Romano, G. and Guerrini, A. (2011), "Measuring and comparing the efficiency of water utility companies: a data envelopment analysis approach", *Utilities Policy*, Vol. 19 No. 3, pp. 202-209.

Saraiya, D. (2005), "The impact of environmental variables in efficiency analysis: a fuzzy clustering-DEA approach", available at: http://scholar.lib.vt.edu/theses/available/etd-08192005-001737/ (accessed November 21, 2013).

Shih, J.-S., Harrington, W., Pizer, W.A. and Gillingham, K. (2004), "Economies of scale and technical efficiency in community water systems", available at: www.rff.org/RFF/documents/RFF-DP-04-15.pdf (accessed November 24, 2013).

Simar, L. and Wilson, P.W. (2011), "Two-stage DEA: caveat emptor", *Journal of Productivity Analysis*, Vol. 36 No. 2, pp. 205-218.

Singh, M.R., Mittal, A.K. and Upadhyay, V. (2011), "Benchmarking of North Indian urban water utilities", *Benchmarking: An International Journal*, Vol. 18 No. 1, pp. 86-106.

Supporting Agency for the Development of Drinking Water Provision System (2013), *Performance of Municipal Water Operators 2012*, SADDWPS, Jakarta.

Tan, J. (2012), "The pitfalls of water privatization: failure and reform in Malaysia", *World Development*, Vol. 40 No. 12, pp. 2552-2563.

Thanassoulis, E. (2000a), "DEA and its use in the regulation of water companies", *European Journal of Operational Research*, Vol. 127 No. 1, pp. 1-13.

Thanassoulis, E. (2000b), "The use of data envelopment analysis in the regulation of UK water utilities: water distribution", *European Journal of Operational Research*, Vol. 26 No. 2, pp. 436-453.

Yasin, M.M. (2002), "The theory and practice of benchmarking: then and now", *Benchmarking: An International Journal*, Vol. 9 No. 3, pp. 217-243.

Zschille, M. and Walter, M. (2010), "Cost efficiency and economies of scale and density in German water distribution", available at: www.diw.de/documents/vortragsdokumente/220/diw_01.c.364475.de/v_2010_zschille_cost_napw.pdf. (accessed November 24, 2013).

About the authors

Dr Andreas Wibowo is a Research Professor at the Indonesia's Ministry of Public Work and a Senior Lecturer at the Graduate Program in Civil Engineering, Parahyangan Catholic University in Bandung, Indonesia. He was formerly an Alexander von Humboldt Postdoctoral Research Fellow at the Bauhaus University in Weimar. Dr Andreas Wibowo is the corresponding author and can be contacted at: andreaswibowo1@yahoo.de

Dr Hans Wilhelm Alfen is a University Professor at the Bauhaus University in Weimar and the Head of Knowledge Centre at Weimar (KC@W), a special purpose vehicle entity as part of an international research network of several top European, African, Asian, and US universities as well as industry partners around the globe.