Socio-Economic Planning Sciences xxx (xxxx) xxx

Contents lists available at ScienceDirect

Socio-Economic Planning Sciences

journal homepage: www.elsevier.com/locate/seps

Environmental sustainability and service quality beyond economic and financial indicators: A performance evaluation of Italian water utilities

Giovanna D'Inverno [a,*], Laura Carosi [b], Giulia Romano [b]

[a] KU Leuven, Faculty of Business and Economics (FEB), Leuven, Belgium
[b] University of Pisa, Department of Economics and Management, Pisa, Italy

ARTICLE INFO

Keywords:
Performance evaluation
Composite indicator
Water utilities
Undesirable indicator
Environmental impact
Sustainable development

ABSTRACT

As water utilities operate as natural monopolists and they provide essential services for human life, their activities are regulated by public authorities. The sustainable use of water resources and a specific attention on social needs should be essential goals for this kind of firms, so that the evaluation of their business should go beyond their profitability and their financial solvency. Keeping pace with the new Circular Economy paradigm and the evolution of the water regulatory framework, in this paper we suggest a global composite indicator apt to evaluate in a novel way the water utilities performance, encompassing financial and economic indexes together with environmental sustainability and service quality measures. To show its empirical implementation we evaluate the performance of Italian water utilities. The operating context is also under scrutiny focusing on specific water utility features such as size, geographical location, degree of diversification and ownership. In this light, operating in the Centre and being large are considered favourable background conditions, while the South and the medium size display a significant unfavourable influence on the water utility performance. Multi-utilities are more advantaged with respect to the mono-utilities and no significant distinction can be made among the different ownership models.

1. Introduction

At the beginning of this century, water has been defined the "blue gold", so to underline the increasing pressure on water resources. On the other hand, as water is an essential requirement, its affordability should be guaranteed to everybody. This is one of the key goals defined by the United Nations (UN) in the 2030 Agenda for Sustainable Development (Goal 6) together with the necessity of a sustainable management and an efficient use of water as a natural resource (Goal 12). In compliance with Goal 6, UN assembly has proclaimed the period from 2018 to 2028 the International Decade for Action "Water for Sustainable Development" (Resolution 71/222). It started on the Water World Day 2018,[1] 22 March 2018, and it will end on World Water Day 2028. With this Resolution, UN has established that "the objectives of the Decade should be a greater focus on the sustainable development and integrated management of water resources for the achievement of social, economic and environmental objectives". The accomplishment of the ambitious objectives of Agenda 2030 can be eased by the good practices of the Circular Economy (CE).[2] According to the European Commission, "CE is a system where the value of products, materials and resources is maintained in the economy for as long as possible, and the generation of waste is minimized" [23]. Therefore, the rational use of resources is one of the cornerstone of the CE approach and it is particularly relevant whenever the availability of natural resources can no longer be taken for granted (e.g. due to water shortages). In the CE context, water and wastewater management occupies a prominent role. In Europe, the integration between the Water Directives and the Circular Economy Action plan requires a greater and greater attention on the water utilities' activities [69]. In many countries such as Italy, England and Wales, the Netherlands and Portugal, these firms operate in a monopolistic context [41,43]. The lack of competition might lead water utilities to provide low quality services with no attention at the environmental issues. Thus, the necessity to define a regulatory framework for the water management in the interest of all the involved stakeholders, namely the consumers that have to be guaranteed quality and affordability of the services, the citizens that are sensitive to

* Corresponding author.
E-mail addresses: giovanna.dinverno@kuleuven.be (G. D'Inverno), laura.carosi@unipi.it (L. Carosi), giulia.romano@unipi.it (G. Romano).

[1] Since 1993, on March 22 the World Water Day underlines the importance of nice practices for sustainable management of freshwater resources. This day is a reminder of the everyday struggle against the water crisis, defined under different aspects.

[2] For an extensive review on the relationship between the Sustainable Development Goals and the Circular Economy practices the reader can see for example [8] and [66].

https://doi.org/10.1016/j.seps.2020.100852
Received 30 October 2019; Received in revised form 27 February 2020; Accepted 8 April 2020
Available online 11 April 2020

environmental issues and the water authorities that have entrusted the water management to the water utilities. While in the past, water utilities were just supposed to ensure service provision and to guarantee their economic and financial sustainability, nowadays they must be socially and environmentally responsible too [40]. This clearly comes to light looking at the evolution of the water regulatory system at both national and international level. It is commonly acknowledged that performance evaluation is a key tool for regulators and that different techniques could be used, from models including key performance indicators, such as the scorecards, to models with overall performance indicators derived from financial ratios or efficiency estimations [30]. The definition of the best practices helps the authorities in defining tariffs and quality standard, in monitoring the service provision and in promoting improvements towards several directions, such as investments, environmental impact, customers' satisfaction, cost efficiency (see for example [18,39,41,43,51,54]). According with the changes on the institutional framework of the water sector and the introduction of the new paradigm of the Circular Economy, different approaches in the water utility performance evaluation have to be adopted. As a firm, a water utility has to be evaluated according to economic profitability and financial solvency criteria, exploiting the data available from the balance sheets, assessing the good or bad management in terms of economic and financial sustainability. Although these criteria have a crucial role in the water utility assessment, they are not representative enough to take the whole picture of the water utility performance. The quality of the service and the environmental sustainability have to be considered too. In this light, water regulators are now linking tariffs to the service quality [44,45]. Specifically, the Italian water regulator has also introduced a new penalty based tariff method for those companies that do not achieve the assigned target in terms of service quality [61].

In the literature, the notion of quality has been declined in different ways, namely the quality of the drinking water supplied, the environmental issues, water losses, unaccounted-for water and customers' perspective. Referring to the latter, several variables have been considered.[3] Among them, it is worth citing the number of the customers' complaints, the values of penalties paid for lack of quality service [46] and the unplanned interruptions [45]. Recently, [61] have used as service quality proxies the promised standard the water utilities have undertaken to comply with, for example in terms of target time to reach all the consumers and to guarantee the water provision. From an environmental point of view, water losses is one of the most used indicators (see for example [31] and [26]). Physical water losses along the distribution net are strictly linked to water crisis risk. Even though a certain level of water losses is considered unavoidable, water utilities should make efforts to reduce them. In drought period, wasting water is socially unacceptable. In the end, customers and more generally citizens will pay for water losses [38]. Moreover, according to the International Water Association [32], the reduction of water losses should be a necessary step in the transition towards the circular economy.[4] Many papers are related to the cost efficiency analysis and include either environmental issues or customer satisfaction indicators. In every case, quality matters [53]: an efficiency analysis with no attention to the quality would penalize those water utilities which make economic efforts in providing a better service. It is worth underlying that there are few papers including together economic, environmental and social issues. In this stream, the analysis of the water sector has taken place in a multicriteria framework. To capture the multifaceted aspects of the performance evaluation, global Composite Indicators (CIs) have been proposed and different methodologies to build them have been suggested (see among others [22,40,43,51,52,54]).[5]

In this paper, we contribute to the literature in two main different directions. First, motivated by the peculiarities of water sector and by the evolution of the regulatory system, we suggest a holistic approach in the utilities performance evaluation to provide decision makers with a comprehensive tool for benchmarking purposes. Specifically, we propose an innovative use of a composite indicator. Second, we show its practical implementation by giving empirical evidence from the Italian water sector. In compliance with part of the literature on performance evaluation (see [20,36,51,70], among others), we opt for a Benefit-of-the-Doubt (BoD) approach in the construction of the composite indicator. The global index is obtained taking a suitable weighted sum of the single performance indicators. In this case, the weights are chosen in the most favourable way for the water utility under evaluation. Accordingly, no firm can complain about being evaluated in an unfair way and about the choice of the weights. We consider a modified version of the BoD model proposed by Zanella et al. [73] to handle the presence of performance indicators with negative values and the so-called "undesirable performance indicators" for which higher values denote lower level of performance (see also [47]). The suggested BoD model allows to include stakeholders' preferences by imposing weight restrictions. Moreover, we consider the robust and conditional version of the outlined composite indicator to mitigate the influence of potential outliers, to directly account for the operating context and to explore its role in the water utility performance [13,15,25,56]. Such a model is applied to 93 Italian water utilities and, as far as we know, this is the first analysis in a global perspective for the Italian case. Beyond the economic and the financial indicators, environmental sustainability and the service quality are taken into account by including water losses, the target time to realize a new connection and the target time to repair ordinary breakdowns. The obtained global index is defined as the Water Utility Performance Composite Indicator (WUP-CI). Finally, the paper gives a contribution to the intense and divisive debate about the influence of the following background condition variables: the size, the ownership, the geographical location and the distinction between mono- and multi-utility. Our analysis shows that operating as a large firm displays a favourable influence on the water utility performance, while the middle size appears the worst one. Referring to the geographical location, being in the Centre is detected as a favourable background condition while being in the South relates negatively to the performance. In line with a part of the recent literature, ownership does not display a significant difference among the management models. As a final evidence, being a multi-utility has a more favourable influence on water utility performance than mono-utility.

The reminder of the paper is organized as follows. In Section 2 the description of the data is presented after a short review of the Italian Water Sector. Data are analysed according to the methodology described in Section 3 and the obtained results are shown and discussed in Section 4. Concluding remarks are the content of Section 5.

2. The water utilities framework

Before going into the methodological and technical details of the Water Utility Performance Composite Indicator (WUP-CI), let us introduce the water utility industry and the Italian context, along with the data.

[3] See for all [61] and the references therein.

[4] IWA suggests a framework towards the CE. In this light it considers three interrelated pathways, namely materials, energy and water. In the latter, water losses is one of the nine crucial issues the Water Utilities should take into account. Regulatory and market levers are also suggested.

[5] For a brief review on the Environmental Composite Indicator construction the reader can refer also to [64].

2.1. The water utility industry

The water industry in Italy has experienced significant changes during the last decades. In the Nineties a broad reform was launched, that fostered an industrialization process and introduced the so-called "Ambito Territoriale Ottimale" (ATO) namely Optimal Territorial Area for the governance of the industry as optimal geographic portions of the regions in accordance with the river basins [5,12,30]. The local supervision of the ATOs was entrusted to local water regulatory authorities. Notwithstanding the reform, the Italian water industry remained highly fragmented, with different management models for thousands of entrusted water operators, with very different features in terms of ownership, size and strategies: from direct management by municipalities or community-owned water supplies, to delegated management to publicly, mixed or privately owned utilities, from very small to big utilities listed on the stock exchange, from monoutilities to multiutilities, from operators that focused only on some water service (for example collection and distribution, sewerage or wastewater treatment) to operators that manage the integrated water services [29].

During the last ten years a broad debate involved the Italian citizens about the best way to manage the water industry; since in 2008 and 2009 the Italian government encouraged through the law the management by mixed or totally private water operators, a much participated referendum was realized in 2011, after which the direct and delegated management by publicly owned firm is still permitted and privatization is no more compulsory [11,29]. After the referendum, still many drawbacks exist and strong need of investments, to overcome the main weaknesses, among which the high leakages, inadequate wastewater treatment, service interruptions and non-drinkable water [29]. So, in 2011 a National Water Authority now called Autorità di Regolazione per Energia Reti e Ambiente (ARERA) was entrusted to supervise and regulate the water industry with the aim to define a homogeneous national framework in terms of tariff method and service contract type and to supervise the role of the local water authorities, directly responsible for supervising the operators that locally manage the water services [61]. Since 2012 ARERA has introduced new tariff methods to encourage investments and improvements in the quality and quantity of water services and to reduce differences among local contexts [59]. Moreover, it introduced standards for the quality of services in terms of contracts and technical requirements and tariffs incentives or penalties for operators that meet or not the targets settled [61]. These standards and targets concern, among others, the reduction of water losses to enforce the efficiency of water use and to reduce environmental impacts and the respect of adequate target time to provide relevant services such as new connection and repair breakdown.

2.2. Data

According with the above mentioned regulatory changes, water utilities have to operate in a new and more demanding institutional framework. As a direct consequence, their assessment should involve several dimensions. Economic and financial indicators have been long considered performance measures of the units under assessment and available in their balance sheets [30,57,60,67]. However, additional indicators are needed when it comes to the evaluation of a unit facing environmental challenges and operating in a regulated market as the water industry is. Beyond the economic and financial status, two more aspects deserve to be accounted for to get a more comprehensive picture of the water utility performance and to support decision making. On the one hand, the environmental aspect needs to be accounted for, given the worldwide commitment and the pressure towards a sustainable management of water resources. On the other hand, the social dimension also has to be considered, so that consumers would not be the only stakeholders bearing the burden of these new challenges resulted in higher bills or in poor service and delay.

In the current study we consider the Italian framework to show the relevance of such a multidimensional and integrated evaluation. The "drought emergency" and the penalties for European laws infringements make Italy an interesting study case to account respectively for environmental and social sustainability issues, in addition to the business indicators commonly used in company performance evaluation. The sample under analysis comprises 93 water utilities for the year 2013.[6] To preserve homogeneity, we focus on those operators that are not municipalities nor public bodies, but only independent companies. Moreover, to provide homogeneity of the data, we consider only those utilities providing all the five main water services (collection, adduction/transportation, distribution of water for civil use, sewerage and wastewater treatment) and reporting all the information for the relevant variables.

Due to the different dimensions involved in the assessment, we assembled an exceptional dataset collecting information from several sources. For the economic and financial items, we use AIDA-Bureau Van Dyck data (https://aida.bvdinfo.com/) and we rely on the most well-known and commonly used indicators. As economic profitability indexes, we observe the Earnings Before Interest, Tax, Depreciation and Amortization (*EBITDA*) *margin* as a proxy for the cash profit. We also consider the Return On Assets (*ROA*) and the Return On Equity (*ROE*) to evaluate how well the company converts respectively assets and investments into net income. Since the net income can be either positive in terms of profits or negative in terms of losses, these two indicators might display negative values that need to be dealt with in the model formulation. As financial indicators, we use both an autonomy index to look at the company's solvency with respect to its own assets (*Financial autonomy*) and with respect to debts from third parties (*Autonomy from third parties*). More in general and beyond the current application, whenever a company's performance is assessed in relation to the information available from balance sheets, economic or financial indicators may admit both positive and negative values whose way to be handled is not trivial.

For the environmental aspect, we include in our analysis the *Percentage of water loss* in the distribution pipes obtained from "Il Portale dell'Acqua" (Italian Water website) available online (http://www.acqua.gov.it/) and realized by the Italian government with data about Italian water operators and municipalities. More precisely, the water loss ratio is obtained as the difference between the water introduced in the network and the water supplied, divided by the water introduced in the network and then transformed in a percentage. The numerator of the ratio represents the leakage from all part of the network, the overflows at the utility's storage tanks and the water thefts. In our dataset we use the average water loss for each water utility obtained as a mean of the water loss in all the municipalities served. This variable has been considered in the water utility performance evaluation in comparatively recent times and included to account for the quality of the service (for a literature review, we refer the interested reader to [61]). However, only in the last years some authors have started distinguishing mostly between two dimensions of the service quality, keeping the environmental sustainability aspect separate from the social one (see [1,45,46]). On the one hand, service quality refers to the environmental impact of the water utility activity, observing for example the quality of water in terms of chemicals in the outgoing water [72], the water losses [26] or the unaccounted-for water [34,53]. On the other hand, service quality comprises the customers perspective on the delivered water services, expressed for example in terms of complaints, unplanned interruptions or time to rectify a sewer blockage. In the present paper, we consider the *Target time to do a new connection (days)* and the *Target time to repair breakdowns (hours)* as service quality indicators to measure the social sustainability of the water utility's

[6] We have chosen data from 2013 since no more recent data are publicly available for the water losses.

Table 1
Descriptive statistics for water utility performance indicators.
Source: Authors' own elaboration based on data from AIDA-Bureau Van Dyck, "Il Portale dell'Acqua" (Italian Water website) and water utilities website.

Water utility performance indicators	Definition	Obs	Mean	St. Dev.	Min	Max
Economic profitability						
ROA	Net Income/Total Assets	93	3.44	3.31	−2.34	20.56
ROE	Net Income/Shareholder's Equity	93	7.55	11.19	−38.12	62.8
EBITDA Margin	Earnings before Interest, Tax, Depreciation, Amortization/Total Revenue	93	21.22	10.64	2.28	45.21
Financial solvency						
Financial autonomy	Shareholder's Equity/Total Assets	93	26.6	18.32	1.95	79.17
Autonomy from third parties	Shareholder's Equity/Total debts	93	0.71	1.03	0.02	5.22
Environmental sustainability						
Water losses (%)		93	0.36	0.11	0.09	0.63
Quality of the service						
Target time to do new connection	Number of days	93	28.54	17.12	7	90
Target time to repair breakdowns	Number of hours	93	2.78	2.63	0.5	12

Table 2
Descriptive statistics for the operating context variables.
Source: Authors' own elaboration based on data from AIDA-Bureau Van Dyck, "Il Portale dell'Acqua" (Italian Water website) and water utilities website.

Background conditions	Obs	Mean	St. Dev.	Min	Q1	Median	Q3	Max
Size								
N of employees	93	274.65	518.98	4	44	127	279	4246
<50 employees	29	31.18%						
50–250 employees	38	40.86%						
>250 employees	26	27.96%						
Ownership								
Publicly owned	66	70.97%						
Mixed-private partnership	27	29.03%						
Geographical location								
North	59	63.44%						
Centre	21	22.58%						
South	13	13.98%						
Diversification								
Multi-Utility	30	32.26%						
Mono-Utility	63	67.74%						

Table 3
Sample means along the performance indicators for *size*.
Source: Authors' own elaboration based on data from AIDA-Bureau Van Dyck, "Il Portale dell'Acqua" (Italian Water website) and water utilities website.

	<50 employees		50–250 employees		>250 employees		Total	
ROA	3.201	(3.090)	2.652	(2.372)	4.858	(4.273)	3.440	(3.312)
ROE	6.118	(7.704)	5.057	(5.192)	12.77	(17.73)	7.545	(11.19)
EBITDA margin	18.60	(10.05)	19.08	(8.753)	27.25	(11.72)	21.22	(10.64)
Financial autonomy	28.46	(20.71)	29.52	(19.21)	20.25	(12.24)	26.60	(18.32)
Autonomy from third parties	0.806	(1.220)	0.870	(1.146)	0.353	(0.276)	0.705	(1.025)
Water losses (%)	0.366	(0.118)	0.351	(0.0976)	0.365	(0.122)	0.360	(0.110)
Target time to do new connection	28.10	(14.66)	25.92	(12.66)	32.85	(23.88)	28.54	(17.12)
Target time to repair breakdowns	2.741	(2.628)	2.592	(2.407)	3.096	(2.997)	2.780	(2.630)
Observations	29		38		26		93	

Note: Standard deviation in parentheses.

activity. These indicators have been already used by Romano et al. [61] with reference to the Italian water utilities operating in two different regions, Tuscany and Veneto, to measure the period for which many customers experienced service interruption inconvenience. The information about these two variables has been retrieved from the service chart, "Carta del Servizio Idrico Integrato", a document drawn up by each utility to report its commitment in providing a certain service standard, approved by the local water authorities and publicly available (in most of the cases, on the water utility's website). For the regulatory period 2016–2019, the Italian Water Authority has conceived an incentive mechanism which links the tariff to the fullfilment of certain standard quality [61]. Referring to the two indicators we consider, the Authority has defined the following targets: 15 days for the time to do a new connection and 3 h for the time to repair breakdowns (https://www.arera.it/it/docs/17/917-17.htm). As the descriptive statistics of Table 1 show, the water utility targets are often way beyond the standard set by the Italian Authority. At this point, it is necessary to reflect on these indicators about environmental and social sustainability aspects and the way they are defined. Differently from the above-mentioned economic and financial indicators, a higher value displayed by these indicators denotes a worse performance. A higher percentage of water losses points at a poorer water utility

Table 4
Sample means along the performance indicators for *ownership*.
Source: Authors' own elaboration based on data from AIDA-Bureau Van Dyck, "Il Portale dell'Acqua" (Italian Water website) and water utilities website.

	Publicly owned		Mixed-Private		Total	
ROA	2.937	(3.364)	4.670	(2.884)	3.440	(3.312)
ROE	6.313	(11.81)	10.56	(9.010)	7.545	(11.19)
EBITDA margin	20.24	(10.02)	23.60	(11.89)	21.22	(10.64)
Financial autonomy	27.57	(20.86)	24.21	(9.538)	26.60	(18.32)
Autonomy from third parties	0.817	(1.189)	0.434	(0.279)	0.705	(1.025)
Water losses (%)	0.348	(0.114)	0.387	(0.0962)	0.360	(0.110)
Target time to do new connection	28.82	(17.89)	27.85	(15.35)	28.54	(17.12)
Target time to repair breakdowns	2.917	(2.785)	2.444	(2.216)	2.780	(2.630)
Observations	66		27		93	

Note: Standard deviation in parentheses.

Table 5
Sample means along the performance indicators for *geographical location*.
Source: Authors' own elaboration based on data from AIDA-Bureau Van Dyck, "Il Portale dell'Acqua" (Italian Water website) and water utilities website.

	North		Centre		South		Total	
ROA	3.339	(3.334)	4.700	(2.821)	1.860	(3.408)	3.440	(3.312)
ROE	7.117	(9.953)	12.38	(9.729)	1.683	(15.64)	7.545	(11.19)
EBITDA margin	20.06	(8.889)	26.26	(13.62)	18.29	(10.79)	21.22	(10.64)
Financial autonomy	28.25	(20.19)	24.40	(12.26)	22.64	(17.71)	26.60	(18.32)
Autonomy from third parties	0.799	(1.129)	0.451	(0.306)	0.692	(1.258)	0.705	(1.025)
Water losses (%)	0.346	(0.107)	0.362	(0.0976)	0.416	(0.131)	0.360	(0.110)
Target time to do new connection	26.75	(12.53)	26	(17.51)	40.77	(28.05)	28.54	(17.12)
Target time to repair breakdowns	2.636	(2.302)	3.167	(3.265)	2.808	(3.045)	2.780	(2.630)
Observations	59		21		13		93	

Note: Standard deviation in parentheses.

Table 6
Sample means along the performance indicators for *diversification*.
Source: Authors' own elaboration based on data from AIDA-Bureau Van Dyck, "Il Portale dell'Acqua" (Italian Water website) and water utilities website.

	Multi-Utility		Mono-Utility		Total	
ROA	3.418	(4.028)	3.450	(2.948)	3.440	(3.312)
ROE	10.72	(12.77)	6.034	(10.12)	7.545	(11.19)
EBITDA margin	17.73	(8.794)	22.87	(11.10)	21.22	(10.64)
Financial autonomy	25.87	(17.39)	26.94	(18.87)	26.60	(18.32)
Autonomy from third parties	0.615	(0.859)	0.748	(1.100)	0.705	(1.025)
Water losses (%)	0.334	(0.0974)	0.372	(0.114)	0.360	(0.110)
Target time to do new connection	27.93	(18.44)	28.83	(16.60)	28.54	(17.12)
Target time to repair breakdowns	2.800	(2.618)	2.770	(2.656)	2.780	(2.630)
Observations	30		63		93	

Note: Standard deviation in parentheses.

performance, harmful to the environment. The same applies for a water utility that reports a longer period of service interruption, detrimental for the consumers. Accordingly, the model formulation has to account for these differences, distinguishing between desirable indicators whose greater values denote a better performance, while the opposite holds for the undesirable indicators. Generally, this reasoning can apply to any indicator, regardless of the classification in the economic, financial, environmental and quality dimensions.

Table 1 shows the descriptive statistics of the water utility performance indicators selected for the current analysis. From an overall look, the multidimensional aspect of the analysis emerges together with the need of accounting for the different preferences of the stakeholders involved in the water industry and in the water utility's activity: the water local authorities, the customers and ultimately the water utilities.

Since the background conditions where a water utility has to operate might influence its performance, we also consider a number of operating context variables, so to ensure a fair benchmarking exercise and to investigate the influence these variables have in the performance assessment. The data choice has been done in compliance with the data availability and the related literature. Specifically, we investigate the relationship between the water utility performance and some aspects that have been subject to reforms in many countries over the last 30 years but whose role is still under debate, namely the size, the ownership and the diversification [29]. In addition to these variables, also the geographical location of the water utilities is taken into account. We collect this information mostly from the websites of the water utilities. The *size* is calculated on the basis of the number of employees following the European Union parameters[7]: utilities with more than 250 employees are large, those with less than 250 and more than 50 employees are medium companies, and those with less than 50 employees are small utilities. As for the *ownership*, we distinguish the mixed-private companies from the public ones [49,57,60]. The degree of *diversification* distinguishes the mono-utilities from the multi-utilities. To conclude, we make a distinction for the water utilities located in the North, the Centre and the South of Italy [58,62]. Table 2 shows the descriptive statistics for the variable size and the distribution of the categorical variables (including the discrete version of size).[8]

Furthermore, to get an intuition of the performance indicators distribution along the operating context variables, we break them down between the different groups. Exploring Table 3 it comes out that

[7] https://ec.europa.eu/eurostat/web/structural-business-statistics/structural-business-statistics/sme.

[8] Due to the high variability of "size" in a continuous scale, in the following analysis we opt for the inclusion of this variable in its discrete version following the legislator definition. The results for the analysis with the continuous variable is available upon request from the authors.

the biggest water utilities exhibit higher means for the profitability indicators and lower means for the financial ones. There is no substantial difference for the water loss indicator, while from the quality side, medium water utilities have higher means. Looking at Table 4 mixed-private companies basically show higher means for the economic indicators and there are not appreciable differences with respect to the other dimensions. Regarding the geographical location, Table 5 displays that water utilities in the South have lower means with respect to all the criteria. Moreover, Table 6 highlights that from an economic perspective four out of five indicators reveal greater means for Mono-Utilities than for Multi-Utilities. However, the situation is overturned with respect to water losses and one of the two quality indicators. The main advantage of using the water utility performance Composite Indicator advocated in this paper is to aggregate all these dimensions in the most favourable way for the units under analysis, weighting the most what they do the best and the least what they do the worst. In this way, none of them can complain of being unfairly assessed, since more importance is assigned to what they perform better and the operating context is taken into account as we are going to explain in the next section.

3. The model

In this section, we explain how to construct the Water Utility Performance Composite Indicator (WUP-CI), accounting for a number of issues raised in the analytical framework outlined so far.

3.1. A directional distance function composite indicator: Undesirable features, negative values and stakeholders' preferences

As we have already pointed out, the assessment of the water utilities' performance should take into account several dimensions, thus the necessity to create a Composite Indicator (CI) to encompass the different criteria under evaluation. In multi-criteria decision analysis (MCDA) different methods have been proposed for the construction of Composite Indicators and then used in different contexts (see for example [51,74]). A large part of the literature on CI is related to the "weighted methods" where a weight is assigned to each criterion and then a weighted sum (or even a product) is determined. In this framework, the weight choice is one of the most crucial aspect and it might be subjective and then questionable. This happens especially when several different stakeholders are involved in the evaluation process and it is difficult to determine, a priori, a hierarchy among the criteria (for an extensive discussion on the weighting of composite indicators, please refer to [28]). The Benefit of the Doubt (BoD) approach appears a suitable methodology to address this issue. According to the BoD, weights are chosen in the most favourable way for the water utility under evaluation,[9] so to give more importance to what it can do the best and low importance to what it does the worst. As the weight determination is data-driven, no evaluated unit can complain for a subjective and penalizing choice. In this perspective, weights are given by granting the water utility the benefit of the doubt [56]. In its first formulation proposed by Cherchye et al. [10], the optimal specific-weights are determined so to make the overall CI as high as possible (an upper bound is fixed as 1). For each unit j_0, the value of CI is obtained by solving the following linear programming problem:

$$\begin{aligned} CI_{j_0} = \max \quad & \sum_{r=1}^{s} y_{rj_0} w_{rj_0} \\ \text{s.t.} \quad & \sum_{r=1}^{s} y_{rj} w_{rj_0} \le 1, \quad \text{for } j = 1, \dots, n, \\ & w_{rj_0} \ge 0, \qquad \text{for } r = 1, \dots, s \end{aligned}$$

where there are S criteria and n units under evaluation, y_{rj_0} is the indicator related to the criterion r for the unit j_0 and w_{rj_0} is the associated weight. The BoD model can be seen as a particular DEA model where several outputs and just a dummy input whose value is fixed at 1 are considered. As it is a DEA-like model, BoD allows to evaluate the performance of water utilities against a frontier consisting of best practice observations. From a practical point of view, the BoD model offers another advantage: the aggregate performance score is unaffected by a rescaling of any of the performance indicator and this allows for the inclusion of ratio data in the analysis. In its first formulation, the BoD model accounts only for criteria that correspond to "desirable" indicators; the higher is the value of the indicator, the better is the performance of the unit under evaluation. In the present analysis, the quality of the service and the environmental sustainability are measured by "undesirable" indicators, that is the higher is their value the worse is the behaviour of the water utilities. To encompass also this second kind of indicators, [73] propose a directional distance version of the BoD model where both the expansion of the desirable indicators and the contraction of undesirable ones are simultaneously considered (see also, [25,36,56]). This is done along the direction of a suitable vector $g = (g_y, g_b)$ where g_y and g_b give the direction for the desirable indicators and the undesirable ones respectively. In this new framework, for each unit, the composite indicator is computed by solving the following maximization problem:

$$\begin{aligned} \max \quad & \beta \\ \text{s.t.} \quad & \sum_{j=1}^{n} b_{kj} \lambda_j \le b_{kj_0} - \beta g_b, \quad \text{for } k = 1, \dots, l, \\ & \sum_{j=1}^{n} y_{rj} \lambda_j \ge y_{rj_0} + \beta g_y, \quad \text{for } r = 1, \dots, s, \\ & \sum_{j=1}^{n} \lambda_j = 1 \\ & \lambda_j \ge 0, \quad \text{for } j = 1, \dots, n \end{aligned} \tag{1}$$

Referring to the unit j, y_{rj} is the r_{th} desirable indicator while b_{kj} is its k_{th} undesirable one. The objective function value at the optimal solution β corresponds to the maximal feasible expansion of the desirable indicators and contraction of the undesirable ones. [73] underline that the directional distance function is comparable to the Shepard's output distance function whenever the directional vector is fixed at the indicators' value of the evaluated unit, i.e. $g = (-g_b, g_y) = (-b_{kj_0}, y_{rj_0})$. The choice of the directional distance function plays a key role in dealing with negative data; actually negative data cannot be straightforwardly included in a DEA model and several approaches have been proposed to overcome this drawback (see for example [2,48,55,68]). Among them [33] and [47] encompass negative data by defining the following directional vector $g = (-g_b, g_y) = (-|b_{kj_0}|, |y_{rj_0}|)$. Their suggestion seems the most preferable one for the current analysis where several water utilities exhibit negative value in the economic profitability indicators.

According to the dual formulation of Problem (1), the weights are chosen in the most favourable way for each unit, without assigning any scale of importance to the single indicators. In some applications, the decision maker's preferences should be taken into account and then indicators cannot be put on equal footing (see for example [21,73]). In the current analysis, the water utilities performance is evaluated paying particular attention on the social and environmental sustainability of their actions. In this light, greater importance to quality and environmental indicators are assigned and this has been done by including assurance region type I (ARI) restriction weights. More precisely, the dual formulation of Problem (1) allows to add constraints regarding the relative importance of each indicator and then the water utility performance can be assessed by solving the following maximization

[9] From now on we will refer to a water utility as the unit under evaluation.

problem [73]:

$$
\begin{aligned}
\beta_{j_0} = \min \quad & -\sum_{r=1}^{s} y_{rj_0}u_r + \sum_{k=1}^{l} b_{kj_0}p_k + v \\
\text{s.t.} \quad & \sum_{r=1}^{s} g_y u_r + \sum_{k=1}^{l} g_b p_k = 1 \\
& -\sum_{r=1}^{s} y_{rj}u_r + \sum_{k=1}^{l} b_{kj}p_k + v \geq 0, \qquad j = 1,\dots,n, \\
& v \in \mathfrak{R}, \\
& u_r \geq 0, \quad \text{for } r = 1,\dots,s, \\
& p_k \geq 0, \quad \text{for } k = 1,\dots,l \\
& \phi_r \leq \frac{u_r \bar{y}_r}{\sum_{r=1}^{s} u_r \bar{y}_r + \sum_{k=1}^{l} p_k \bar{b}_k} \leq \psi_r \quad \text{for } r = 1,\dots,s \\
& \phi_k \leq \frac{p_k \bar{b}_k}{\sum_{r=1}^{s} u_r \bar{y}_r + \sum_{k=1}^{l} p_k \bar{b}_k} \leq \psi_k \quad \text{for } k = 1,\dots,l
\end{aligned}
\tag{2}
$$

where $\bar{y}_r$ and $\bar{b}_k$ are related to the indicators of the ideal average unit. The last two constraints require that the percentage contribution of any indicators must vary within a specific range, namely $[\phi_r, \psi_r]$ for the desirable indicators $r = 1,\dots,s$ and $[\phi_k, \psi_k]$ for the undesirable ones $k = 1,\dots,I$. In this way, the weight constraints are the same for every evaluated unit. Once the maximum value for Problem (2) is determined, the performance measure is given by $CI_{j_0} = 1/(1+\beta)$. The best performing units have a score equal to one, while for the others the composite indicator scores vary from zero to one.

3.2. A robust and conditional approach of the directional distance function composite indicator: Outlying observations and operating context

In the estimation of Problem (2), there are two more aspects that need to be taken into account. First, we have to tackle the presence of outlying observations (if any) in the sample under analysis, to avoid downwardly biased estimates in the sample under analysis [25]. The literature has proposed a number of approaches to handle this issue. Among others, outlier detection procedures have been suggested to identify potential outlying observations, such as atypical observations and/or measurement errors, and to remove them from the sample. However, this could result in the removal of some observations that in their atypical magnitude might carry very useful and relevant information [17]. Neglecting this piece of information would work against the aim of performing a fair and comprehensive benchmarking exercise. Therefore, we opt for a different solution. Following the seminal papers by Cazals et al. [9] and adapted in the composite indicator framework as proposed by De Witte and Rogge [19], Rogge et al. [56] and Lavigne et al. [36] among others, we adopt a robust approach. The main intuition behind is to mitigate the influence of outlying observations by executing a Monte-Carlo simulation. Specifically, we draw a high number of B times (e.g. $B = 2000$) and with replacement $m < n$ observations. For each of the B draws, we compute a water utility performance composite indicator $CI^{b,m}$ for the m-size sample. Then, we compute the robust composite indicator CI^{m} as the arithmetic average of the computed $CI^{b,m}$. Due to the sub-sampling, the unit under analysis might not be part of the reference set. In this case, unlike the non-robust CI score, the robust CI score can take values larger than one. Accordingly, the evaluated water utility is deemed super-performing and interpreted as doing better than the average m randomly drawn utilities in the reference sample [20].

The second aspect concerns the context where the unit, in our case the water utility, has to operate in. It might be the case that working in a certain environment or institutional setting rather than another one might favour or hamper the achievement of a certain level of performance. For this reason, we should account for these background variables that are not under the direct control of the managers, but that still might exert an influence on the distribution of the performance scores and the attainable set. The literature presents several ways to investigate the role of these characteristics and different solutions have been adopted in the water utility framework as well (see for example [24,42,71]). In this context, we consider a conditional approach, following the path traced by Daraio and Simar [13,15] and adjusted to the composite indicator context as in [20,25,36] among others. The choice of this approach leads to two main advantages. First, it avoids imposing the so-called "separability condition" by including in one-stage the contextual variables in the estimation of the score (see for all [13,14,16,56]). Differently from the unconditional case, where each unit is equally likely to be drawn in the sampling with replacement, in this case the m units are drawn B times with replacement but with a particular probability estimated by means of a kernel function on z background variables. By performing the sampling in this way, the unit under analysis will be compared with units operating in an environment more similar to its own. At this point, the water utility is fairly assessed not only because of the favourable weighting system, but also because of the units it is compared with. The robust conditional water utility composite indicator $CI^{m,z}$ is obtained as the arithmetic average of the conditional composite indicator $CI^{b,m,z}$ computed B times. In this case, a $CI^{m,z}$ value larger than one suggests that the water utility under analysis is doing better than the average m utilities in the reference sample facing a similar background. As a second benefit from performing a conditional analysis, we can explore the potential influence of the background characteristics besides the benchmarking exercise. A non-parametric statistical inference on the relationship between the estimated composite indicator scores and the background variables under scrutiny can be done. For an intuitive interpretation of the results (see for all [56]), we non-parametrically regress the ratio between the robust unconditional CI^{m} and conditional $CI^{m,z}$ scores with the operating context variables z and the statistical significance can be obtained (see for all [16,37]). In this way, a positive slope in the partial regression plot denotes a favourable influence of the considered variable on the water utility performance level and the opposite holds for a negative slope. We refer the interested reader to all the above mentioned references for a more extensive and technical discussion.

4. Empirical application

In this section, we present the results on the comprehensive performance level assessment of 93 Italian water utilities in 2013. We obtain this evidence by using the model specification described in the previous section. Specifically, we estimate the water utility performance score along different dimensions, namely the economic and the financial ones as well as the environmental and service quality ones. Before exploring the main findings, two considerations need to be made and kept in mind while interpreting the results.

As mentioned in Section 3.1, imposing the weights restricts the role that each indicator can play in the overall performance assessment and the extent to which it can contribute to the water utility performance score. In the context of the water industry, the role of the regulator is to mediate the interests of all the involved stakeholders, including the companies, the customers and ultimately the environment itself. The measures adopted by the national authority in the last decades keep track of the order of priorities as well. For example, penalties have been envisaged for those water utilities that commit European laws infringement in order to safeguard the environment. Similarly, tariff penalties have been laid down for those companies that provide a customer disservice. Following this rationale, incentive measures have been also taken to reward the companies that comply with the law and achieve good service targets [61]. The present analysis reflects the priorities of the national regulator on the different dimensions under scrutiny to fairly account for this multidimensional framework. Specifically, referring to Problem (2) in Section 3, lower (i.e. ϕ_r) and upper (i.e. ψ_r) bounds have been set for each indicator, so to give the following scale of importance among the analysed dimensions: 1. environmental, 2. service quality, 3. economic sustainability, 4.

Table 7
Water Utility Performance Composite Indicator (WUP CI) estimates for different model specifications.
Source: Authors' own elaboration.

	Mean	St. Dev.	Min	Q1	Median	Q3	Max
Unconditional WUP CI	0.6143	0.1313	0.4041	0.5051	0.6079	0.6962	1.0000
Robust unconditional WUP CI	0.6946	0.1514	0.4686	0.5659	0.6982	0.8097	1.0883
Robust conditional WUP CI	0.7394	0.1569	0.4344	0.6058	0.7331	0.8547	1.0451

The conditional model accounts for *Size*, *Ownership*, *Location* and *Diversification*.

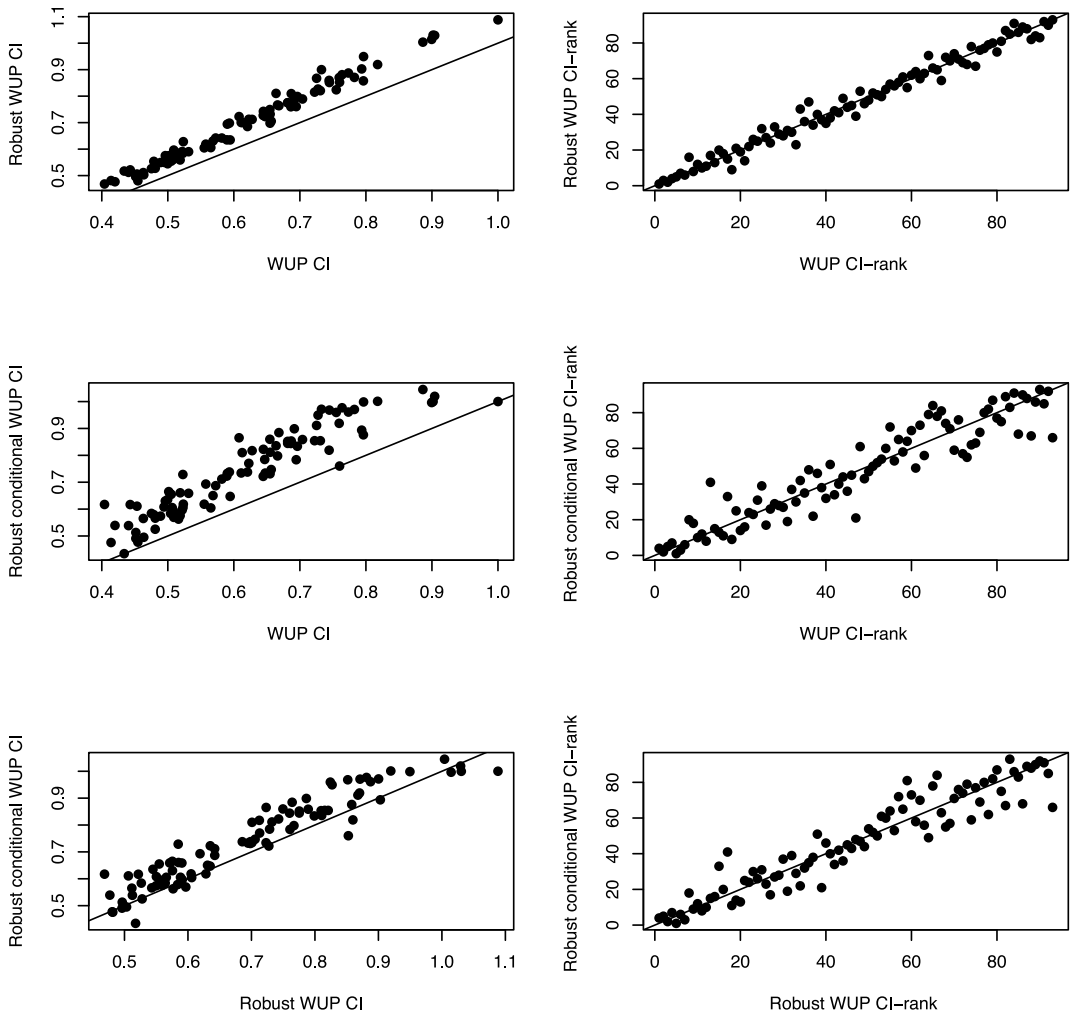

Fig. 1. Scatter plots of the Water Utility Performance Composite Indicator (WUP CI) scores and ranks for different model specifications.
Source: Authors' own elaboration.

financial sustainability. Service quality indicators share the same lower and upper bounds. ROA and Financial autonomy have the highest upper bound among the economic and the financial indicators since they are strictly linked to the invested capital. Being water utilities capital intensive firms, particular attention should be given to the investment side. On the contrary, dividend distribution is not very frequent in the water sector and Water Authority encourages firms in reinvesting their earnings to improve the quality of the provided services [58]. To balance the interests of all the involved stakeholders, the chosen upper bounds impose that the environmental and quality dimensions give a relatively higher contribution with respect to the financial and the economic one. If the evaluation had been performed reflecting the priorities of the companies, the weights assigned to each indicator would have most likely been different and the same reasoning applies if the customers priorities had been considered.[10] Moreover, the performance scores have been estimated by accounting for some operating context variables deemed important in the literature, namely the size, the ownership, the geographical location and the diversification. In Section 2.2 few patterns have already emerged by looking at the descriptive statics of the performance indicators grouped by

[10] We have also estimated two different versions of the composite indicator following opposite extreme cases with respect to this baseline model. Moving away from the balancing goal of the regulator, one extreme scenario is to place very high priority on the environmental and quality service indicators, while the opposite extreme on the economic and financial indicators. In either cases, the change in the importance of each indicator weights does play a role on the Water Utility Performance Composite Indicator scores. We refer the reader to Appendix for an extensive discussion of the main findings.

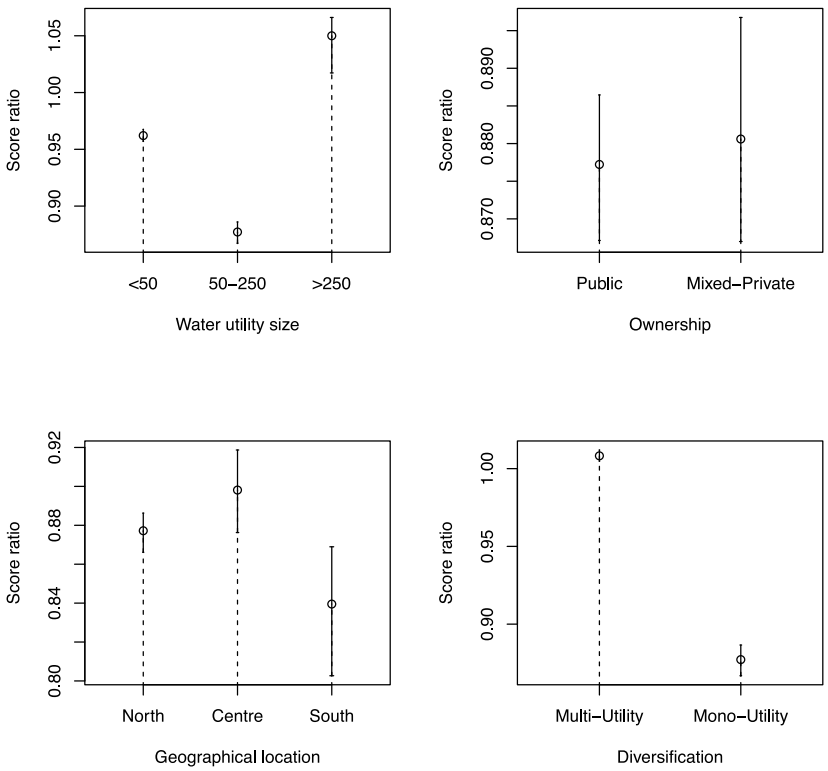

Fig. 2. Visualization of the partial regression plots with bias-corrected bootstrapped nonparametric confidence intervals for the operating context variables. *Source:* Authors' own elaboration.

different categories. In the following we show the potential of including these variables all together, so to get a broader picture and to capture synergies among these characteristics.

Table 7 shows the results of the estimated robust conditional water utility performance composite indicator (WUP CI). For completeness and comparison purposes, we report also the non-robust and the robust unconditional version of the advocated composite indicator.[11] The robust conditional approach together with the non-robust and the robust unconditional one all offer a more integrated picture of the units under analysis. In fact, the non-robust case can be seen as a reference starting point. The robust approach gives the idea of the presence or not of atypical observations that might affect the findings of the non-robust analysis and it mitigates their influence. The conditional analysis allows to explore the role of external variables on the overall evaluation, both in the score estimation process and in the comparison between the robust unconditional scores and the conditional ones. Taken as a whole, we get a more exhaustive idea of the framework under investigation while accounting for different aspects. Going to the computed scores, the mean of the non-robust unconditional score is equal to 0.6143, suggesting a sizeable room for performance improvement by looking at the best practices detected in the sample, namely 38.57% obtained as 1-0.6143. Once the robust approach is considered to take into account the possible presence of atypical observations, the score slightly increases to 0.6946. Then, the potential room for improvement suggested when both the atypical observations and the operating context are accounted for amounts to 26.06%. This number should represent an indicative starting point to boost its own level of performance, while looking at the best practices and complying with the priorities of the national regulator. Before moving the discussion on the influence that the context has with respect to the detected level of performance, the descriptive statistics displays at least a couple of more very informative aspects. First, the minimum values and the first quartile denote the presence of very poorly performing water utilities. Second, at the other extreme, the maximum values of the robust unconditional and conditional scores suggest the presence of at least one super-performing water utility.

Fig. 1 provides a graphical comparison of the water utility performance composite indicator in its non-robust, robust unconditional and robust conditional version (for other applications, see also [19,56,70], among others). The comparison between the robust and non-robust CI scores in terms of ranking shows the observations (the water utilities) along the 45-degree line. This implies a negligible role of the robust approach in the sample under analysis. Moreover, all the robust scores are consistently higher than in the non-robust case. By comparing the conditional and the unconditional CI scores, we can appreciate the importance of taking into account the operating context to ensure a fairer benchmarking exercise. When looking at the distribution according to the rank, we see that many observations lie below the 45-degree line. This suggests that these water utilities benefit from the conditional assessment and specifically from accounting for an unfavourable context that might hamper the achievement of higher level of performance.

We explore the potential relationship among the water utility performance and some operating context variables usually considered in the literature. Specifically, we compute the ratio of the unconditional over the conditional water utility performance composite indicators to draw some considerations. We caution the reader that this is based on correlational evidence, so we refrain from giving a causal interpretation of the findings. We report the partial regression plots in Fig. 2 to explore whether background conditions have either a favourable or an unfavourable influence on the performance assessment [56,65,70].[12] To focus the discussion of the results, we report only the evidence from

[11] The estimates have been obtained by choosing m=20 after a sensitivity analysis and B=2000. For more technical details we refer to Section 3.2.

[12] The statistical tests indicate that the operating context variables are significantly related to the ratio between robust unconditional and conditional scores at all conventional levels (see for all [16,37]).

the complete conditional model, accounting for all the four context variables. The inclusion of fewer ones delivers similar results available upon request from the authors.

Regarding the *size*, we can observe that large utilities have a positive and statistically significant relationship with the water utility performance, as suggested by the small and not overlapping confidence intervals. On the contrary, operating as a medium utility has the most unfavourable influence on the water utility performance. This evidence is in line with previous findings on the Italian water industry framework (see for example [29]), suggesting the idea that water utilities can perform better when big players or, to some extent, when they remain small local firms [63]. However, it is worth pointing out that the role of the size as an operating context variable is still controversial. Previous studies have not unequivocally identified the "optimal dimension" for a water utility and the findings strongly differ from one country to another [6,7,29]. Concerning the *ownership*, we do not find major differences in the way they correlate with the ratio between robust unconditional and conditional scores and the confidence intervals are quite broad. This confirms what have been found already in the literature (see for example [49,50] and the literature reviews by Guerrini et al. [30] and Berg and Marques [4]). One possible explanation can be rooted in the fact that this topic had been long debated in the last decades and many reforms have been enacted in this regard, giving rise to a more balanced coexistence of these different organizations. Moreover, privatization processes do not necessarily lead to improvements in efficiency and quality of service provision [3,35] as well as in tariff increases/decreases [27,59]. The *geographical location* presents a pattern consistently detected in other papers, that is, the South shows a significant unfavourable influence when assessing the performance of the water utilities compared to the Northern region and even more to the Centre. Most of this evidence can be interpreted in the light of two main aspects. First, the morphological characteristics and the climate cause long drought periods and hence water utilities face many difficulties in guaranteeing the water provision. Second, there are better infrastructure and more investments in the Centre and in the North compared to the southern regions. The southern mains are generally old and so, paradoxically, the South is both characterized by water scarcity and by the highest rate of water loss in Europe [29]. About the *diversification*, being a mono-utility shows a significant unfavourable influence on the performance of a water utility. Positive correlation with multi-utility might come from higher level of performance obtained by taking advantage of potential synergies arising from operating in different sectors, such as gas and electricity.

5. Conclusion

The strategic role of water provision and sanitation for sustainable development entails the necessity to regulate the water utilities' activity. The regulatory evolution, the new objectives and commitments of national and local Water Authorities, the new paradigm of the Circular Economy impose to evaluate the water sector from a multidimensional perspective. Therefore, beyond the traditional economic and financial criteria, the detection of best practices and virtuous behaviours should be carried out by including also the environmental and the quality issues of the service. In this light, we have proposed an innovative use of a non-parametric composite indicator, here labelled Water Utility Performance Composite Indicator (WUP-CI), to aggregate the following dimensions: the economic profitability, the financial solvency, the water losses and the customer satisfaction. The composite indicator has been obtained by running a directional distance BoD-model in its three versions: the non-robust, the robust and the robust conditional one. The model turns out to be suitable for dealing with crucial issues of the analysis, such as the presence of indicators with negative values, the presence of "undesirable indicators", the necessity of including in the analysis both the stakeholders' preferences and the background characteristics. The WUP-CI has then been used to evaluate the performance of 93 Italian water utilities that provide the integrated water service, including distribution, sewerage and wastewater activities. The presence of weight restrictions in the model has allowed to give more emphasis on the environmental and the quality dimension, rather than on the economic and the financial one, in accordance with the Service of General Economic Interest (SGEI) nature of the urban water service. Moreover, alternative scenarios for stakeholders' preferences have been also considered, showing that different priorities do play a role in the overall performance assessment. The empirical analysis showed that the Italian water industry needs policy and strategies to improve the performance of water utilities. Decision makers can get valuable information looking at the best practices that emerge from the analysis. Specifically, evidence can suggest where to intervene and to which extent to boost the performance of those water utilities with unsatisfactory results. Moreover, the analysis can contribute to the long-standing debate on what the relevant operating context variables are and how they influence the water utilities' performance. With reference to the size, our results show a positive relationship with this factor. On the contrary, operating in the South of Italy and being a mono-utility company can be considered as unfavourable conditions as compared to the background of the water utilities belonging to the other groups. Therefore, specific policies to support southern regions and to foster the performance of mono-utilities should be defined to overcome their gap. With respect to the ownership, no significant performance differences emerge between the totally publicly owned operators and the partially publicly owned ones. This finding is in line with the strand of the literature demonstrating that privatization processes should not be enforced, leaving the management model choice to decision makers on the basis of context specificities.

CRediT authorship contribution statement

Giovanna D'Inverno: Conceptualization, Methodology, Software, Data curation, Investigation, Formal analysis, Writing - original draft, Writing - review & editing. **Laura Carosi:** Conceptualization, Methodology, Formal analysis, Writing - original draft, Writing - review & editing. **Giulia Romano:** Conceptualization, Data curation, Investigation, Writing - original draft, Writing - review & editing.

Acknowledgements

We would like to thank the anonymous referees, Kristof De Witte and the participants at EWEPA 2017 in London, DEA 2017 in Prague, IWA Conference in Livorno and AMASES XLI in Cagliari for their valuable comments and suggestions on a previous version of the paper. Giovanna D'Inverno acknowledges financial support from Research Foundation – Flanders, FWO, Belgium (Postdoctoral Fellowship 12U0219N), Laura Carosi and Giulia Romano acknowledge financial support from University of Pisa, Italy (PRA 2018 Project).

Appendix. Alternative scenarios for stakeholders' preferences

As already mentioned above, the nature of the water utility industry makes the environmental aspects and the service quality as well as the economic and financial indicators essential to provide a comprehensive evaluation of the water utilities' performance. However, the importance that can be assigned to each indicator and to each area does play a role in the overall assessment. In the following we consider different scenarios where different priorities are assigned to the sub-indicators in the construction of the water utility performance composite indicator, showing how the change in the importance of each indicator weights affects the final ranking and the role of the contextual variables on the Water Utility Performance Composite Indicator (WUP-CI).

Specifically, we have estimated two different versions of the composite indicator following opposite extreme cases with respect to the

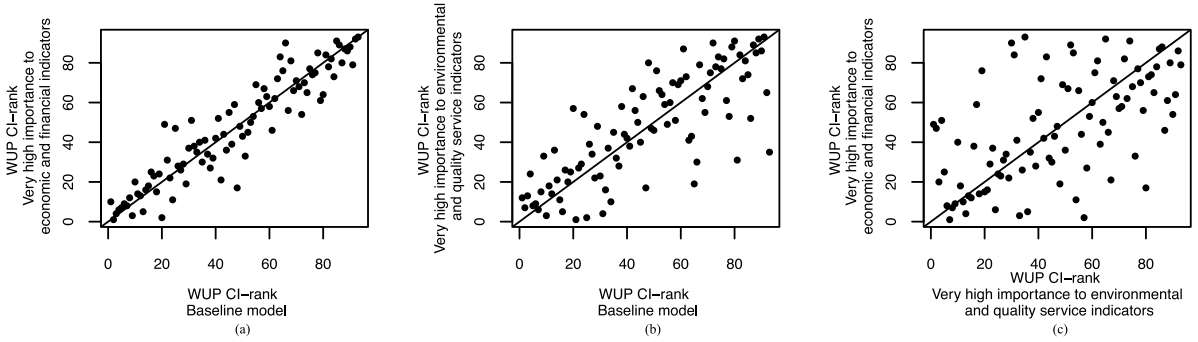

Fig. A.1. Scatter plots of the Water Utility Performance Composite Indicator (WUP CI) ranks for different scenarios.
Source: Authors' own elaboration.

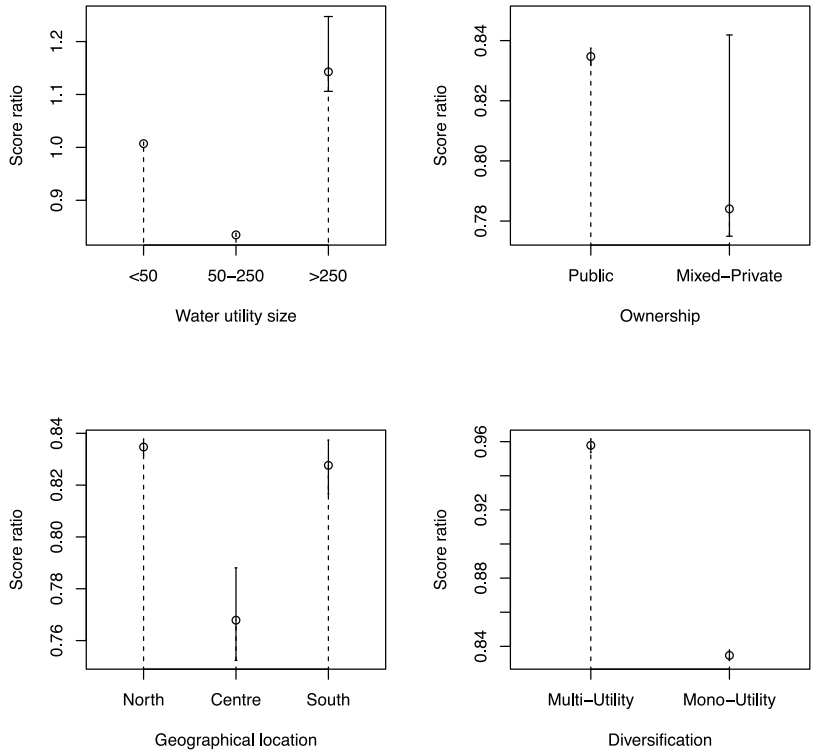

Fig. A.2. Visualization of the partial regression plots with bias-corrected bootstrapped nonparametric confidence intervals for the operating context variables. *Very high importance to economic and financial indicators*.
Source: Authors' own elaboration.

baseline model as presented in Section 4. The baseline composite indicator is constructed so to reflect the role of the regulator in mediating the interests of all the involved stakeholders (e.g., the companies, the customers and the environment), looking for a balance between the increasing need of the legislator on environmental and service quality issues and the market power of the companies themselves. Moving away from the balancing goal of the regulator, one extreme scenario is to place very high priority on the environmental and quality service indicators, while the opposite extreme on the economic and financial indicators.

In Fig. A.1 we present three scatter plots of the Water Utility Performance Composite Indicator (WUP CI) ranks for different scenarios, to show how the change in the importance of each indicator can affect the final ranking. Fig. A.1(a) provides a graphical comparison of the WUP CI ranks between the baseline model and the scenario where very high importance has been assigned to economic and financial indicators. Fig. A.1(b) compares the WUP CI ranks of the baseline model and the scenario where very high importance has been assigned to environmental and service quality indicators. In both cases, many observations lie below or above the 45-degree line. This points at the fact that the overall water utility performance evaluation might be subject to different outcomes depending on the interests pursued by the analysis. In the baseline model, the detected best practices should guide the regulator in identifying the most critical areas and in setting the system of penalties and tariffs accordingly from a policy perspective. Reasonably, there will be water utilities that turn out to be penalized by this kind of assessment, either because they devote most of their attention in the economic and financial dimensions – one extreme – or because very committed to the environmental cause or to the customer satisfaction – the other extreme (we refer to [62], for a critical discussion of the Italian framework in this regard). In the former case the water utilities overlook their impact on the environment and offer a lower service quality to the customers. In the latter case, the water utilities neglect an economic and financial sustainability need required by their capital intensive nature, in favour of a too much stronger attention to the environment and the quality that eventually backfires on the customers by imposing higher tariffs. Taking this argument to the extreme, Fig. A.1(c) provides a graphical comparison of the WUP CI

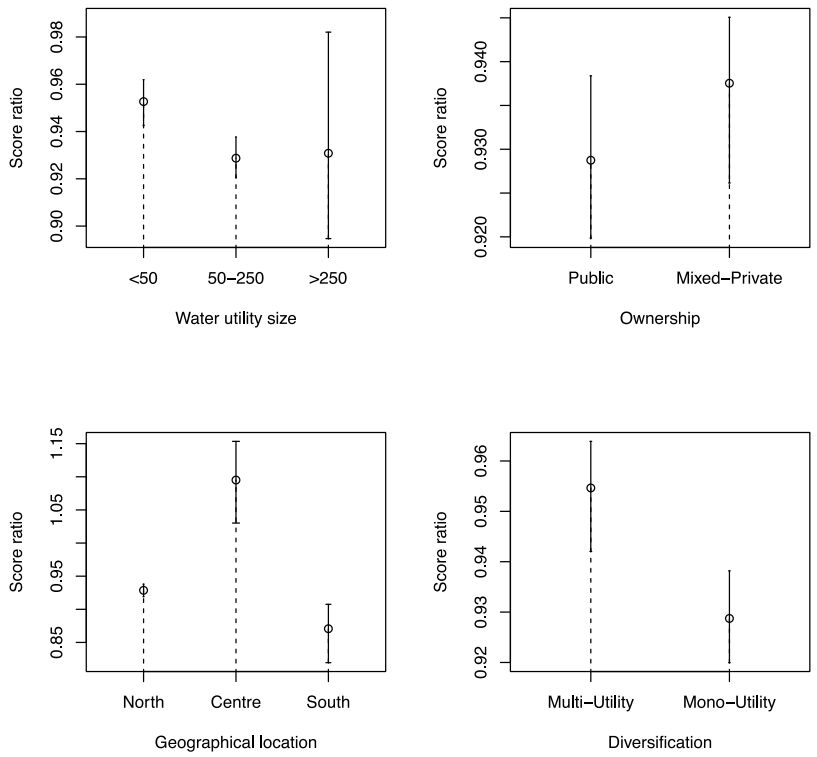

Fig. A.3. Visualization of the partial regression plots with bias-corrected bootstrapped nonparametric confidence intervals for the operating context variables. *Very high importance to environmental and service quality indicators*.
Source: Authors' own elaboration.

ranks between these two extreme scenarios. In this case the impact in the ranking and in the detected best practices is sizeable. Similar patterns emerge when the non-robust and the robust unconditional version of the Water Utility Performance Composite Indicator is considered.

Following the same rationale, we question whether the role of the structural aspects considered in this analysis might change depending on these two extreme evaluation scenarios. Figs. A.2 and A.3 report the partial regression plots to explore the role of the contextual variables on the Water Utility Performance Composite Indicator (WUP-CI) estimated giving very high importance to the economic and financial aspects, or to the environmental and service quality ones. The main considerations that can be drawn concern mostly the geographical location and the size of the water utilities. In the scenario where very high importance is given to the economic and financial indicators, the Centre displays a negative relationship with the water utility performance compared to the North and the South. This evidence is slightly overturned in the baseline scenario and very pronounced once very high importance is attributed to the environmental and service quality indicators. A plausible explanation can be found in the strong commitment undertaken by the water companies located in the Centre of Italy to sustainable water and information campaigns to support it, confirming the evidence provided by previous studies on the Italian water industry framework (see for example [62]). As for the size, while in the scenario with greater importance on the economic and financial indicators and in the baseline model being a large firm (>250 employees) shows a significant favourable influence on the performance of the water utility and being a medium firm (50–250 employees) displays the most unfavourable influence, this is not the case anymore when giving very high importance to the environment and service quality. In fact, smaller water utilities could be more keen to accept customers and citizens' requests in terms of service provided and attention to avoid unnecessary environmental impact and waste of precious resources as drinkable water. For this reason, including environmental and service quality indicators in performance assessment helps to improve the evaluation of smaller utilities with respect to medium and big ones. The overlapping confidence intervals prevent us from inferring any statistically significant difference among size classes. Concerning the degree of diversification, similar patterns emerge compared to the baseline model, that is, being a multi-utility has a positive relationship with the utility performance. In both scenarios, there are no statistically significant differences among different types of ownership, confirming the evidence arising from the baseline scenario. This result confirms that publicly owned management models (direct or delegated) are able to obtain the same results in terms of performance of totally or partly privatized models when also environmental and service quality is included in the analysis. This result fosters previous findings [61] showing that publicly owned utilities performed only slightly worse than public–private partnerships when quality issues are included in the efficiency assessment. Thus, privatization should be evaluated with caution when only economic and financial indicators are included in the analysis since other relevant aspects of the mission of public utilities (in terms of environmental sustainability and service quality to citizens) are omitted [53].

Overall, we can conclude that giving either more importance to the economic and financial or to the environmental and the quality aspects does matter and play a role.

References

[1] Ananda J, Pawsey N. Benchmarking service quality in the urban water industry. J Prod Anal 2019;51(1):55–72.
[2] Aparicio J, Pastor JT, Vidal F. The directional distance function and the translation invariance property. Omega 2016;58:1–3.
[3] Bakker K. Privatizing water: Governance failure and the world's urban water crisis. Cornell University Press; 2010.
[4] Berg S, Marques R. Quantitative studies of water and sanitation utilities: A benchmarking literature survey. Water Policy 2011;13(5):591–606.
[5] Carrozza C. Italian water services reform from 1994 to 2008: Decisional rounds and local modes of governance. Water Policy 2011;13(6):751–68.

[6] Carvalho P, Marques RC. Computing economies of vertical integration, economies of scope and economies of scale using partial frontier nonparametric methods. European J Oper Res 2014;234(1):292–307.

[7] Carvalho P, Marques RC. Estimating size and scope economies in the Portuguese water sector using the Bayesian stochastic frontier analysis. Sci Total Environ 2016;544:574–86.

[8] Cavaleiro de Ferreira A, Fuso-Nerini F. A framework for implementing and tracking circular economy in cities: The case of porto. Sustainability 2019;11(6):1813.

[9] Cazals C, Florens J-P, Simar L. Nonparametric frontier estimation: A robust approach. J Econometrics 2002;106(1):1–25.

[10] Cherchye L, Moesen W, Rogge N, Puyenbroeck TV. An introduction to "benefit of the doubt" composite indicators. Social Indic Res 2007;82(1):111–45.

[11] Ferreira da Cruz N, Marques RC, Romano G, Guerrini A. Measuring the efficiency of water utilities: A cross-national comparison between Portugal and Italy. Water Policy 2012;14(5):841–53.

[12] Danesi L, Passarelli M, Peruzzi P. Water services reform in Italy: its impacts on regulation, investment and affordability. Water Policy 2007;9(1):33–54.

[13] Daraio C, Simar L. Introducing environmental variables in nonparametric frontier models: A probabilistic approach. J Prod Anal 2005;24(1):93–121.

[14] Daraio C, Simar L. Advanced robust and nonparametric methods in efficiency analysis: Methodology and applications. Springer Science & Business Media; 2007.

[15] Daraio C, Simar L. Conditional nonparametric frontier models for convex and nonconvex technologies: A unifying approach. J Prod Anal 2007;28(1–2):13–32.

[16] De Witte K, Kortelainen M. What explains the performance of students in a heterogeneous environment? Conditional efficiency estimation with continuous and discrete environmental variables. Appl Econ 2013;45(17):2401–12.

[17] De Witte K, Marques RC. Influential observations in frontier models, a robust non-oriented approach to the water sector. Ann Oper Res 2010;181(1):377–92.

[18] De Witte K, Marques RC. Gaming in a benchmarking environment. A non-parametric analysis of benchmarking in the water sector. Water Policy 2011;14(1):45–66.

[19] De Witte K, Rogge N. Accounting for exogenous influences in performance evaluations of teachers. Econ Educ Rev 2011;30(4):641–53.

[20] De Witte K, Schiltz F. Measuring and explaining organizational effectiveness of school districts: Evidence from a robust and conditional Benefit-of-the-Doubt approach. European J Oper Res 2018;267(3):1172–81.

[21] D'Inverno G, De Witte K. Service level provision in municipalities: A flexible directional distance composite indicator. European J Oper Res 2020. http://dx.doi.org/10.1016/j.ejor.2020.04.012, in press.

[22] Duarte AA, Rodrigues GMdC, Ramos RA. A global service quality index to evaluate the performance and sustainability in water supply utilities. WSEAS Trans Environ Dev 2009;5(12):759–69.

[23] European Commission. Closing the loop — An EU action plan for the circular economy. Communication no. 614. 2015, https://ec.europa.eu/environment/circular-economy/.

[24] Fuentes R, Torregrosa T, Ballenilla E. Conditional order-m efficiency of wastewater treatment plants: The role of environmental factors. Water 2015;7(10):5503–24.

[25] Fusco E, Vidoli F, Rogge N. Spatial directional robust Benefit of the Doubt approach in presence of undesirable output: An application to Italian waste sector. Omega 2019.

[26] Garcia S, Thomas A. The structure of municipal water supply costs: Application to a panel of French local communities. J Prod Anal 2001;16(1):5–29.

[27] González-Gómez F, García-Rubio MA. Prices and ownership in the water urban supply: A critical review. Urban Water J 2018;15(3):259–68.

[28] Greco S, Ishizaka A, Tasiou M, Torrisi G. On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. Soc Indic Res 2019;141(1):61–94.

[29] Guerrini A, Romano G. Water management in Italy: Governance, performance, and sustainability. Springer; 2014.

[30] Guerrini A, Romano G, Campedelli B. Factors affecting the performance of water utility companies. Int J Public Sect Manag 2011;24(6):543–66.

[31] Hernández-Sancho F, Molinos-Senante M, Sala-Garrido R, Saz-Salazar D, et al. Tariffs and efficient performance by water suppliers: An empirical approach. Water Policy 2012;14(5):854–64.

[32] IWA. Water utility pathways in a circular economy. IWA, London: International Water Association; 2016.

[33] Kerstens K, Van de Woestyne I. Negative data in DEA: A simple proportional distance function approach. J Oper Res Soc 2011;62(7):1413–9.

[34] Kirkpatrick C, Parker D, Zhang Y-F. An empirical analysis of state and private-sector provision of water services in africa. World Bank Econ Rev 2006;20(1):143–63.

[35] Kishimoto S, Lobina E, Petitjean O. Our public water future: The global experience with remunicipalisation. Transnational Institute (TNI)/Public Services International Research Unit (PSIRU)/Multinationals Observatory/Municipal Services Project (MSP)/European Federation of Public Service Unions (EPSU); 2015.

[36] Lavigne C, Jaeger SD, Rogge N. Identifying the most relevant peers for benchmarking waste management performance: A conditional directional distance Benefit-of-the-Doubt approach. Waste Manag 2019;89:418–29.

[37] Li Q, Racine JS. Nonparametric econometrics: Theory and practice. Princeton University Press; 2007.

[38] Lombardi G, Stefani G, Paci A, Becagli C, Miliacca M, Gastaldi M, et al. The sustainability of the Italian water sector: An empirical analysis by DEA. J Cleaner Prod 2019;227:1035–43.

[39] Marques RC. A yardstick competition model for Portuguese water and sewerage services regulation. Util Policy 2006;14(3):175–84, doi:10.1016/j.jup.2006.03.004.

[40] Marques RC, da Cruz NF, Pires J. Measuring the sustainability of urban water services. Environ Sci Policy 2015;54:142–51.

[41] Marques RC, Simões P, Pires JS, et al. Performance benchmarking in utility regulation: The worldwide experience. Pol J Environ Stud 2011;20(1):125–32.

[42] Molinos-Senante M, Gémar G, Gómez T, Caballero R, Sala-Garrido R. Eco-efficiency assessment of wastewater treatment plants using a weighted Russell directional distance model. J Cleaner Prod 2016;137:1066–75.

[43] Molinos-Senante M, Marques RC, Pérez F, Gómez T, Sala-Garrido R, Caballero R. Assessing the sustainability of water companies: A synthetic indicator approach. Ecol Indic 2016;61:577–87.

[44] Molinos-Senante M, Maziotis A, Mocholí-Arce M, Sala-Garrido R. Accounting for service quality to customers in the efficiency of water companies: Evidence from England and Wales. Water Policy 2016;18(2):513–32.

[45] Molinos-Senante M, Maziotis A, Sala-Garrido R. Estimating the cost of improving service quality in water supply: A shadow price approach for England and Wales. Sci Total Environ 2016;539:470–7.

[46] Molinos-Senante M, Sala-Garrido R, Lafuente M. The role of environmental variables on the efficiency of water and sewerage companies: A case study of Chile. Environ Sci Pollut Res 2015;22(13):10242–53.

[47] Oliveira R, Zanella A, Camanho AS. The assessment of corporate social responsibility: The construction of an industry ranking and identification of potential for improvement. European J Oper Res 2019;278(2):498–513.

[48] Pastor JT, Ruiz JL. Variables with negative values in DEA. In: Zhu J, Cook WD, editors. Modeling data irregularities and structural complexities in data envelopment analysis. Spr inger; 2007, p. 63–84.

[49] Peda P, Grossi G, Liik M. Do ownership and size affect the performance of water utilities? Evidence from Estonian municipalities. J Manag Gov 2013;17(2):237–59.

[50] Pérard E. Water supply: Public or private? an approach based on cost of funds, transaction costs, efficiency and political costs. Policy Soc 2009;27(3):193–219.

[51] Pérez F, Delgado-Antequera L, Gómez T. A two-phase method to assess the sustainability of water companies. Energies 2019;12(13):2638.

[52] Pérez F, Molinos-Senante M, Gómez T, Caballero R, Sala-Garrido R. Dynamic goal programming synthetic indicator: An application for water companies sustainability assessment. Urban Water J 2018;15(6):592–600.

[53] Picazo-Tadeo AJ, Sáez-Fernández FJ, González-Gómez F. Does service quality matter in measuring the performance of water utilities? Util Policy 2008;16(1):30–8.

[54] Pinto F, Costa A, Figueira J, Marques R. The quality of service: An overall performance assessment for water utilities. Omega 2017;69:115–25.

[55] Portela MS, Thanassoulis E, Simpson G. Negative data in DEA: A directional distance approach applied to bank branches. J Oper Res Soc 2004;55(10):1111–21.

[56] Rogge N, De Jaeger S, Lavigne C. Waste performance of NUTS 2-regions in the EU: A conditional directional distance Benefit-of-the-Doubt model. Ecol Econom 2017;139:19–32.

[57] Romano G, Guerrini A. The effects of ownership, board size and board composition on the performance of Italian water utilities. Util Policy 2014;31:18–28.

[58] Romano G, Guerrini A. Paying returns to shareholders of water utilities: Evidence from Italy. Sustainability 2019;11(7):2033.

[59] Romano G, Guerrini A, Campedelli B. The new italian water tariff method: A launching point for novel infrastructures or a backwards step? Util Policy 2015;34:45–53.

[60] Romano G, Guerrini A, Vernizzi S. Ownership, investment policies and funding choices of Italian water utilities: An empirical analysis. Water Resour Manag 2013;27(9):3409–19.

[61] Romano G, Molinos-Senante M, Guerrini A. Water utility efficiency assessment in Italy by accounting for service quality: An empirical investigation. Util Policy 2017;45:97–108.

[62] Romano G, Salvati N, Guerrini A. Factors affecting water utility companies? Decision to promote the reduction of household water consumption. Water Resour Manag 2014;28(15):5491–505.
[63] Romano G, Salvati N, Guerrini A. Governance, strategy and efficiency of water utilities: The Italian case. Water Policy 2018;20(1):109–26.
[64] Sabia G, De Gisi S, Farina R. Implementing a composite indicator approach for prioritizing activated sludge-based wastewater treatment plants at large spatial scale. Ecol Indic 2016;71:1–18.
[65] Schiltz F, De Witte K, Mazrekaj D. Managerial efficiency and efficiency differentials in adult education: A conditional and bias-corrected efficiency analysis. Ann Oper Res 2019;1–18.
[66] Schroeder P, Anggraeni K, Weber U. The relevance of circular economy practices to the sustainable development goals. J Ind Ecol 2019;23(1):77–95.
[67] Shaoul J. A critical financial analysis of the performance of privatised insustries: The case of the water industry in England and Wales. Crit Perspect Account 1997;8(5):479–505.
[68] Silva MC, Camanho AS, Barbosa F. Benchmarking of secondary schools based on students' results in higher education. Omega 2019;102119.
[69] Smol M, Adam C, Preisner M. Circular economy model framework in the European water and wastewater sector. J Mater Cycles Waste Manag 2020;1–16.
[70] Verschelde M, Rogge N. An environment-adjusted evaluation of citizen satisfaction with local police effectiveness: Evidence from a conditional Data Envelopment Analysis approach. European J Oper Res 2012;223(1):214–25.
[71] Villegas A, Molinos-Senante M, Maziotis A. Impact of environmental variables on the efficiency of water companies in England and Wales: A double-bootstrap approach. Environ Sci Pollut Res 2019;1–12.
[72] Woodbury K, Dollery B. Efficiency measurement in Australian local government: The case of New South Wales municipal water services. Rev Policy Res 2004;21(5):615–36.
[73] Zanella A, Camanho AS, Dias TG. Undesirable outputs and weighting schemes in composite indicators based on data envelopment analysis. European J Oper Res 2015;245(2):517–30.
[74] Zarghami M, Szidarovszky F. Multicriteria analysis: Applications to water and environment management. Springer Science & Business Media; 2011.

Giovanna D'Inverno is FWO Postdoctoral Fellow at the Faculty of Economics and Business, KU Leuven, in the 'Leuven Economics of Education Research' center (Belgium). She received her joint Ph.D. in Economics at IMT School for Advanced Studies Lucca (Italy) and in Business Economics at KU Leuven (Belgium) in 2018. Her research interests focus on quantitative methods and policy evaluation methods for public economic issues, in particular local government expenditure and school spending.

Laura Carosi is Full Professor in Applied Mathematics for Economics at the Department of Economics and Management, University of Pisa (Pisa). She has her Ph.D. degree in Mathematics for Economic Decisions, University of Florence (Italy). Currently her research interests are fractional programming, scalar and vector optimization, generalized convexity, data envelopment analysis, general equilibrium models with incomplete markets.

Giulia Romano is Associate Professor in Business Administration at the Department of Economics and Management, University of Pisa (Pisa). She has her Ph.D. in Business economics, University of Pisa (Italy). Her research activity focuses mainly on public service provision, performance evaluation and corporate governance, in particular with reference to public utilities.