Durham University Business School

Electricity Distribution in Poland and Great Britain:
A Comparative Efficiency Analysis 2013-2016

Z0961626

16 March 2018

Electricity Distribution in Poland and Great Britain: A Comparative Efficiency Analysis 2013-2016

Executive Summary

This paper reviews the recent experience of the Polish electricity distribution sector under incentive regulation. In 2015, the Polish Energy Regulation Office implemented a framework regulating the individual cost of capital based, among others, on benchmarking of own performance. This paper uses benchmarking in the international context to compare the largest four Polish and fourteen British DSOs, using two-stage semi-parametric bootstrapped DEA and cost SFA. In the first stage, bias corrected efficiencies are calculated. In the second stage, the efficiencies are regressed on an environmental variable representing the country. The paper finds consistent results across the two models, and in comparison to external models based on Bayesian SFA. The applied methodology favours a parsimonious model with relatively few explanatory variables, which account for most of the variations in efficiency scores. In particular, quality of service indicators are accounted for by other factors, such as the network size. The Polish companies are underperforming in comparison to their British counterparts. The DEA model finds that Polish companies are on average 50% less efficient, whereas the SFA model estimates that the value of TOTEX will increases by 180% when the company is Polish. The paper makes several observations on the differences in the regulatory regimes and makes policy recommendations with regards to transparency.

Declaration

This dissertation is the result of my own work. Material from the published or unpublished work of others, which is referred to in the dissertation, is credited to the author in question in the text. The dissertation is 11,000 words in length.

Acknowledgements

I would like to thank my supervisor, Professor Tooraj Jamasb, for his support, kindness and clear advice, without which this dissertation would not have been possible.

Also, I would like to thank Mr Kamil Makieła for providing his vital input and extensive explanations.

My deepest gratitude goes to my parents for their constant support and love.

Table of Contents

List of Abbreviations

Capex – Capital Expenditure

CRS – Constant Returns to Scale

DEA – Data Envelopment Analysis

DRS – Decreasing Returns to Scale

DSO – Distribution System Operator

NRA – National Regional Authority

Ofgem – Office of Gas and Electricity Markets

Opex – Operating expenditure

SFA – Stochastic Frontier Analysis

Totex – Total Expenditure

UK – United Kingdom

URE – Urząd Regulacji Energii; Energy Regulation Office

VRS – Variable Returns to Scale

1. Introduction

1.1. Motivating remarks

Power outages in Poland are some of the longest in Europe, with rural areas being worst affected. In 2015, planned and unplanned SAIDI (average outage duration for each customer served) for the Polish DSOs was 394 minutes, whereas for the UK it was 60.39 minutes (CEER, 2015). The outages in Poland last approximately six times longer than in Western Europe. Up to 40 % inhabitants of Poland live in rural areas, where outages happen most often, thereby affecting those living in greatest poverty. The responsibility for the management and control of the electricity distribution network belongs to regional Distribution System Operators (DSOs)[1], which operate as regulated natural monopolies. They are responsible for providing equal access to network to all consumers.

A significant investment burden lies on the DSOs, which are required to implement Smart Grid solutions and charging stations for electric cars in urban areas. At the same time, the distributed generation[2] reshapes how electricity networks are operated and managed. This creates an environment in which the DSOs are exposed to significant pressure for efficiency and innovation.

In 2015, the Polish Energy Regulation Office (URE) introduced a new benchmarking framework with an aim of reducing the time and frequency of outages in the system, as well as decreasing the connection time. The new framework sets out to reduce the values of SAIDI and SAIFI by 50 % by 2020. URE consulted the British Ofgem when devising their long-term strategy and the new regulatory framework. The motivation of this study is therefore to compare the two countries, with regards to their regimes and performance of individual companies and see whether the new regulation improved the quality-incorporated efficiency of the Polish companies.

1.2. Introduction

Electric power distribution is the final stage of the delivery of electric power. Electricity distribution is commonly operated as a regulated natural monopoly. Modern utility regulation uses benchmarking of companies to decrease the information asymmetry that exists between the regulator and the companies and sometimes to directly affect the behaviour of firms. Efficiency analysis requires comparability of firms. Therefore it is natural to restrict

[1] Referred to as Distribution Network Operators (DNOs) in Great Britain.
[2] Defined as production of electricity near to the place of consumption.

the sample to domestic utilities. However, an international comparison may be a valuable tool, given a small sample of domestic companies, to evaluate the performance of national utilities within the context of international best practice.

The performance of British DSOs under incentive regulation has been the subject of productivity and efficiency analysis in the past, using both national and international samples. However, no such comparisons were performed for Poland after the incorporation of quality benchmarking by the Polish regulator. This paper tries to address this gap and examines the efficiency of electricity distribution utilities in Poland in comparison to the British DSOs, using panel data from 2013 to 2016. It employs a two-stage Data Envelopment Analysis (DEA) with a bootstrap estimation, in which bias-corrected efficiency estimates are calculated in the first stage and then regressed on external environmental variables in the second stage. Furthermore, it verifies the results using the Stochastic Frontier Analysis (SFA). The results of both models are then assessed using an external model, based on the Bayesian SFA. The aim of this paper is to: i) suggest and implement an approach for international comparisons, which provides consistent results, ii) to quantify the performance gap between the two countries, and iii) make recommendations on how to improve the accountability standards, hoping that it will foster academic and business research in the field.

This paper is structured as follows: Chapter 2 outlines the literature on the public utility regulation and productivity analysis, with a focus on electricity distribution. Chapter 3 describes the current regulatory frameworks in Poland and Great Britain in the light of the regulatory economics. Chapter 4 describes the methodology used for performing the efficiency analysis. Chapters 5 and 6 report on the data and results. Chapter 7 concludes and makes policy recommendations and suggests ideas for further research.

2. Literature review

This chapter presents the relevant literature in the field of natural monopoly regulation, explains the different regulatory regimes, and highlights the issues involved with the quality of service and international benchmarking.

2.1. Public Utility Regulation

Electricity distribution is described as a natural monopoly, due to significant sunk costs and economies of scale which imply low marginal costs of provision (Kahn, 1971). The infrastructure investments include substations, transformers, overhead lines, poles and meters. These investments tend to be lumpy and have long life spans. Capacity costs are the driving cost factor in electricity distribution, while services they provide are non-storable. Utility regulation then aims to strike a balance between optimal capacity expansion, which requires cost-coverage, and optimal capacity utilisation, which requires fluctuating prices (Vogelsang, 2002).

In the absence of regulatory control, energy distribution firms will face low internal incentives for efficiency and may engage in rent-seeking behaviour, which leads to allocative inefficiency and suboptimal provision of services. An unconstrained monopolist would restrict output to maximise profits, thus leading to a transfer of welfare, otherwise enjoyed by consumers (in a competitive setting), to producers. Furthermore, the lack of competitive threat would lead to productive inefficiency. Therefore, electricity distribution companies are subject to utility regulation.

The regulator is assumed to be maximising the aggregate consumer welfare in the presence of two constraints: maintaining the long-term viability of the utility and imperfect information (Joskow and Schmalensee,1986). The primary goal of public utility regulation is to encourage the regulated firms to produce output efficiently at the cost (meeting the demand) and service quality (reliability) dimensions (Joskow, 2008). At the same time, the economic regulation of monopoly networks requires the determination of revenue sufficient to cover costs, and a return on equity high enough to attract capital from investors. According to the European Commission Directive (2009), the role of the National Energy Offices is to:

> *"ensure that transmission and distribution system operators are granted appropriate incentive, over both the short and long term, to increase efficiencies, foster market integration and security of supply and support the related research activities".*

The task of the regulatory agency is to ensure that this maximum revenue does not provide an excessive rate of return or reimburse the company for an excessive level of costs (Shuttleworth, 2005). However, the economic theory does not give a precise definition of what is an “adequate”, and what is an “excessive” rate of return, thus lending itself to ambiguity in implementation by individual regulators.

Modern regulatory literature[3] views the problem as a game between a principal (the regulator) and agents (the regulated firms). In practice, the regulator faces imperfect information about the relevant cost, quality, and demand attributes faced by the regulated firm. Joskow and Schmalensee (1986) notice that the asymmetric information, on the side of the regulator, occurs despite the availability of large amounts of information on the accounting costs of regulated firms, published in the forms of reports and consultations. Full information would require the knowledge of the underlying marginal costs, necessary for cost-setting, and available opportunities and threats. Firms have an information advantage and thus discretion in choosing the input bundle and production method. In theory, firms attempt to maximise their profit, taking the regulatory mechanism as a given. In reality, the relationship between the regulator and the utility is closer than with the consumers. The firms may enjoy a degree of determination with regards to the shape of the regulatory mechanism and the choice of parameters. Furthermore, firms may decide to exploit strategically their information advantage to increase their profits or reach some managerial goals.

Ideally designed regulation would pass the economies of natural monopoly and network reliability on to customers while providing shareholders with a fair return. Frontier Economics (2010) points out that, given the uncertainty of technological innovation, successful regulation would focus on encouraging frontier technology solutions, and not only keeping customer costs to a minimum. Profit motive may not be in line with promoting the quality of service and “(…) as a non-tradable product, the resultant quality level tends to deviate from the socio-economic optimum” (Giannakis et al., 2003)

Regulators must make decisions regarding the efficient minimum size of the monopolies when they assign a specific territorial jurisdiction. Economies of scale and scope may exist with respect to prices and quality of service. Natural monopolies are the most efficient when the

---

[3] Vogelsang (2002) refers to this as the Bayesian regulatory mechanism. In the Bayesian paradigm, the regulator’s lack of information is described by subjective prior probabilities, that the regulator holds about the parameters of the optimisation problem.

entire demand can be satisfied at lowest cost by one firm rather than by two, or more[4] (Posner, 1968). Filipini (1998) finds a minimal optimal size of regional distribution companies and supports franchised monopolies, rather than side-by-side competition. Kwoka (2005) argued that smaller utilities may supply higher quality, as they have easier access to local market and customer-specific information such as demand characteristics and the specific technical conditions. Also, it has been argued that proximity to customers leads to higher quality of service (and lower losses). Empirical studies do not support this hypothesis. Growitsch et al. (2005) show, using a sample of European utilities, that countries with many smaller utilities (e.g. Norway) tend to be significantly less efficient than large utilities from, for example, Italy or the UK. Yatchew (2000), looking at municipal electric utilities in Canada, finds evidence of increasing returns to scale in small companies, and constant or decreasing returns in larger firms. At the same time, decision has to be made regarding the ownership structure, be it private or public. Kumbhakar and Hjalmarsson (1996) find evidence that privately owned Swedish DSOs are relatively more efficient, than their publicly owned counterparts.

2.2. Regulatory regimes

In this section, we can explain the two most commonly used regimes for regulation of natural monopolies, which vary with regards to its incentive power:

- Rate of return regulation,
- Fixed price (*revenue*) schemes (*price-cap*, *revenue cap*, *RPI-X*).

Also, we highlight the concept of "yardstick competition": a regulatory mechanism, in which the performance of various agents is compared and rewards or penalties are assigned to agents on basis of their relative performance.

Both fixed price schemes and yardstick competition are described as incentive regulation, the purpose of which is to induce the utility to achieve desired goals, whilst leaving the discretion of the method to the companies. In doing so, it emulates the discipline of the market upon the firm. However, the outcome of incetive regulation is often ambigious, and regualtors tend to adopt a combination of different mechanisms to achieve their objectives (Jamasb and Poudineh, 2013).

2.2.1. Rate of Return Regulation

[4] i.e. have a subadditive cost function. Subadditivity has been demonstrated theoretically and empirically in the infrastructure networks (see Gilsdorf, 1995 for US electricity utilities; Salvanes and Tjøtta, 1998 for Norway).

Rate of Return (RoR) regulation schemes are a low-powered option to determine the financial return of an industry. The primary goal is providing necessary compensation to investors. In performing this form of regulation, the regulator determines the reimbursement (regulated income) in a given year (t) for a company *k*, by setting the appropriate amount for the company's rate base $(K^k(t))$, cost of capital with an appropriate mark-up $(r+\delta)$, operating expenses (C^k_{OpEx}), and depreciation $(D^k(t))$ (see (1)). Capital investment deemed prudent is covered, but the OPEX is usually capped proportionally to the regulatory asset base. The burden of proof often rests on the regulator, and it involves considerable regulatory administration efforts in order to avoid imprudent or unreasonable operating expenditures and investments to enter the compensation and rate base.

$$R^k(t) = C^k_{OpEx}(t) + D^k(t) + (r+\delta)K^k(t) \tag{1}$$

Since energy market liberalisation, regulators have tended to move away from cost-recovery regulation towards incentive-based regulation. RoR encourages utilities to invest in capacity at levels that exceed the socially optimal level (Averch and Johnson, 1962; Zajac, 1970, 1972; Gal-Or and Spiro, 1992). In case that the RoR exceeds the cost of capital firms substitute capital for labour to increase profit, which leads to a high capital-labour ratio and therefore allocative inefficiency. The RoR regulation is used to stimulate investments in networks and technologies because it transfers a part of the risk from the investor to the society. In the past, such regulation was practical in the early phase of electrification, as it induces large capital-intensive investments. It failed, however to promote productive efficiency and cost-saving investments (Cambini and Rondi, 2010).

2.2.2. *Fixed price regimes*

The RPI-X price cap, popularised in the 1983 Littlechild Report (Stern, 2003), and its revenue cap variant are examples of high-powered regimes. Price cap regulation was first implemented in the post-privatisation regulation of British Telecom and was since adopted in many countries. See (2) for a generic representation. Under price (or revenue) cap, cost allowances $(C^k(0))$ are set in advance for a fixed period (five years in the UK), with adjustment for predicted productivity development per year x (i.e. inflation) and individual requirements x^k (the X-factor)[5] to reflect the level of historical costs and thereby the need to catch-up to best practice.

[5] The offsetting X-factor is the sum of: the difference in the total factor productivity growth rates in the regulated sector and rest of the economy, and the difference in the input price growth rates

$$R^k(t) = C^k(0)(1 - x - x^k)^t,\ t = 1, \dots, T \tag{2}$$

A regulated network may retain financial benefits if it outperforms the underlying assumptions of the allowed revenue calculation. Similarly, if a company underperforms, it must bear at least part of the associated cost (Ofgem, 2009). To maximise the profit, the company needs to minimise the cost. Liston (1993) showed that the fixed income induces cost efficiency by the agent's cost minimisation. Theoretically, the specified rate of price decline is decoupled from the realised performance and provides a strong incentive for cost reductions. The key issue is the level of the X-factor: if it is too low, then the rents will be excessive; if it is too high, then the firm's financial viability is threatened.

The benefits of price cap regulation include providing companies with incentives to improve efficiency, decreasing the effects of cost information asymmetries between firms and regulators, lowering regulatory costs and reducing the incentives to over-invest in the capital and cross-subsidise relative to the rate of return regulation. In a dynamic setting, a price cap subjects businesses to more risk (measured by beta), thus creating incentives for internal efficiency (Alexander and Irwin, 1996). Balazs (2009) argues that an incentive regulation (together with an independent regulator) introduces a more coherent regulatory framework which favours investment, whereas under an ad hoc, discretionary RoR regulation, under-investment can occur. Regulatory opportunism introduces uncertainty into the relationship and is, therefore, a concern for regulated companies. Shifts in the regulatory frameworks which allow for the use of cost disallowance instruments will decrease the propensity to invest (Gal-Or and Spiro, 1992; Lyon and Mayo, 2005). Also, Besanko and Spulber (1992) show, under the RoR regulation, that the regulators' lack of commitment can lead to underinvestment. Cambini and Rondi (2010) highlight, using a sample of European energy utilities, that the investment rates in cost reducing projects, such as the Smart Grid solutions, are greater under incentive regulation than under rate of return regulation. Furthermore, they find that investment rates are highly sensitive to the X-factor.

However, the price cap is associated with several theoretical and practical problems. Agrell and Grifell-Tatjé (2016) prove, using a game-theoretic setting with a probability of judicial repeal of a high-powered regulation, that regulated firms may maintain cost-inefficiency even under slumping profitability. If the implemented model is not economically sound and compatible, the firms no longer trust in the long-term commitment from the regulator, and

between the economy and the regulated sector (Bernstein and Sappington, 1998). This could be extended to include service quality variations.

their investment strategy deviates from cost efficiency. Therefore, "the implementation of price cap mechanisms is more complicated and their efficiency properties more difficult to evaluate than it is often implied" (Joskow, 2008).

2.2.3. *Yardstick competition*

Yardstick competition arises when a regulator uses outside information from several comparable firms to determine the incentives for each firm (Schleifer, 1985). The idea is to mimic the conditions of a market by using observations to estimate a real cost function in each period. The permitted revenue for DSO would be set ex-post and determined by the costs in the same period of other firms operating under similar conditions. The aim of such exercise is that the outcomes of perfect competition are replicated in a regulated natural monopoly context.

There is a risk of collusion, especially if there are few operators. Also, a set of comparators with correlated operating conditions must be established. Finally, the firms should all not be far off from the best practice, as the underperformers are likely to go bankrupt. Weitzman (1980) and Freixas et al. (1985) describe the "ratchet effect", which might arise when firms, subject to performance targets, restrict their efficiency because they rationally anticipate regulators to respond by "raising the bar".

2.3. Quality of Service

The quality of power distribution can be divided into the quality of customer services (measurements, responsiveness), reliability (outage time and frequency) and voltage quality (the characteristics of the voltage wave).

Ajodhia and Hakvoort (2005) divide quality regulation instruments into three categories: (1) indirect instruments, (2) standards, and (3) incentive schemes. Indirect instruments are those actions designed to improve the negotiation position and information of consumers. Examples include obligatory publishing of information on companies' performance. Standards are minimal levels of expected quality, below which consumers are entitled for a compensation from the DSO. In an incentive scheme, a company is compared to a target, internal or external, which results in either a fine or a reward.

The challenge for the regulator is to define a country-specific, socially desirable level of the quality of service, in terms of continuity of supply, rather than merely maximising the quality level. The regulator needs to equate the marginal cost of improving service quality with the marginal reward from that level of service based on the customers' preferences. Alternatively, quality is optimal for a monopolist if the profits are maximized (Ajodhia and Hakvoort, 2005). Then, the task is to internalise the rewards of the quality provision into the utility's profit function, thus aligning the incentives (Growitsch et al, 2010). Higher quality is often more expensive to provide, and therefore results in higher prices for consumers. Spence (1975) showed that an unregulated monopolist might either under or over provide quality, according to the preferences of the marginal consumer relative to the average. Profitability may also be positively linked to the quality of service, in certain dimensions. Ogden (1997) identifies prompt bill payment, lower costs of handling complaints and reputation as three incentives to improve quality in the water industry.

The effects of different regimes have ambiguous effects on the quality of service. RoR regulation is likely to increase incentives for the over-provision of quality if its provision is capital intensive and would boost the rate base (Sherman, 1989 in Price et al., 2002). Ter-Martirosyan (2003) analysed the impact of incentive regulation on the duration and frequency of electric outages for a panel of 78 US utilities. The study finds that incentive regulation may be associated with an increase in the average duration of electric outages. However, the implementation of specific quality benchmarks can reduce the average outage duration.

Cost efficiency and quality of service seem to be two distinct targets. Giannakis et al. (2005) found, by using DEA to study the service quality of electricity distribution utilities in the UK, that cost-efficient companies had not necessarily shown high service quality. They also concluded that integrating service quality in regulatory benchmarking was preferable to cost-only approaches. Accordingly, some economic regulators set both price caps and quality standards. Generally, since privatisation, the permitted price levels have deflated, while the quality standards have been raised (Price et al., 2002).

2.4. Benchmarking

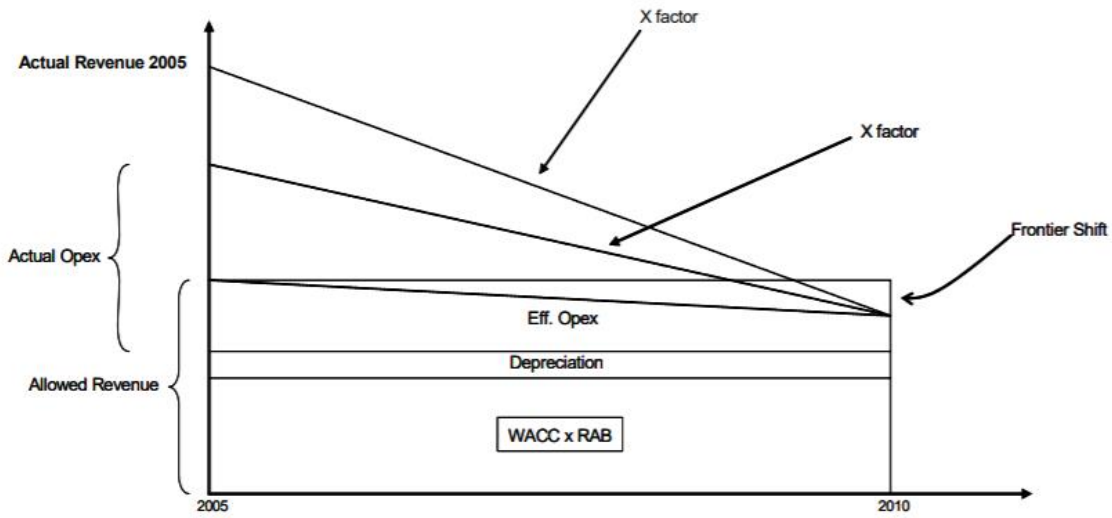
2.4.1. Aims and Uses

Benchmarking can be broadly defined as a comparison of some measure of realised efficiency and productivity performance against a benchmark performance (Jamasb and Pollitt, 2000). Furthermore, benchmarking can be performed in comparison to other firms, for example as a means of imposing "yardstick competition", or to own historical performance. Benchmarking allows comparative regulation and, in the first case, uses outside information besides the information revealed by an individual company. Hence, it helps to eliminate or reduce the firm's asymmetric information advantage with regards to its inputs (operational and capital costs) and demand.

To guide the choice of individual and general requirements in the incentive regimes, the regulator performs systematic benchmarking[6]. The regulator rewards agents based on their relative performance and therefore generates incentives for promoting efficiency. This is achieved by constructing a benchmark, based on average or frontier performance of the industry. The general requirement x can be set using Malmquist analysis (panel data over the years before the regulatory period), whereas individual requirements x^k are typically linked to the individual inefficiencies in the period prior to the regulatory review (Agrell and Bogetoft, 2016). Furthermore, the required catch-up varies across countries[7]. Figure 1 illustrates the Ofgem regulatory period 2005-2010, where firms are allowed to recover their capital cost and are required reach the adequate levels of OPEX and allowed revenue at the end of the period.

Figure 1

OPEX benchmarking and determination of allowed revenues and X-factors



Source: Jamasb and Politt, 2007

Likewise, the quality of service performance can be evaluated using benchmarking. Figure 2 shows an example of how service interruptions feed into the benchmarking process and thus affect the utility's overall profits[8].

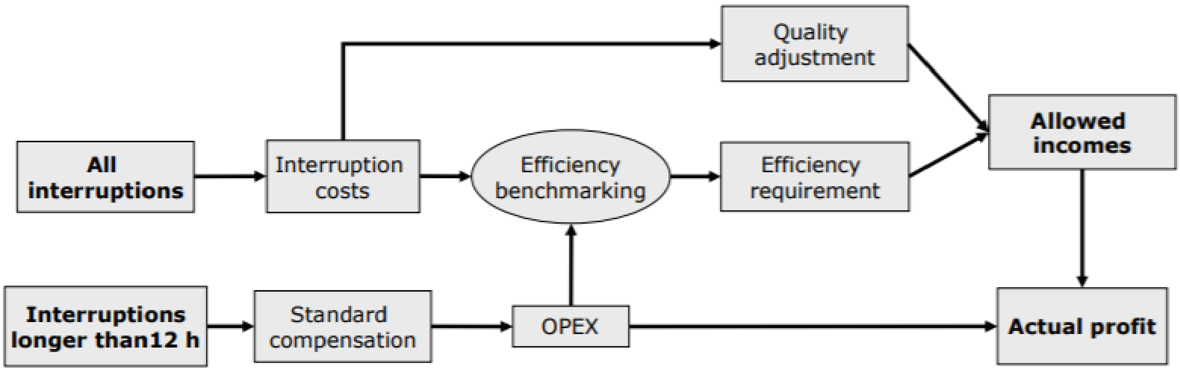
[6] Typically, every 5 years, at the outset of the regulatory period. The yardstick regime requires the same analysis as a fixed price regime, but the benchmarking takes place more frequently (e.g. annually).

[7] Some countries require the DSOs to catch-up very quickly (Bogetoft and Otto, 2010). Denmark required inefficiencies to be eliminated in 1 year, the Netherlands in 3-5 years; Germany in 10 years. Reducing the performance gap may take longer than initially anticipated (see Pollitt, 2005 for case of Ofgem).

[8] For example, Ofgem uses individual benchmark to set targets for planned interruptions. These are calculated as the average number of interruptions and minutes lost over the previous three years for each DSO. For unplanned interruptions, the targets are calculated by benchmarking across the DSOs and looking at each DSO's historical performance (Ofgem, 2017).

Figure 2

Example of the effects of interruptions on the DSO's allowed income



Source: Honkapuro, 2008, p.164

Benchmarking is not without its issues, both in theory and practice. Shuttleworth (2005) lists a number of problems related to benchmarking: (i) difficulty in choosing the correct distribution of error and inefficiency term, (ii) inconsistent results produced by the different models used, (iii) low ratio of observations on companies and their costs relative to the number of potentially relevant explanatory variables, (iv) negative bias towards companies less involved with the regulator, (v) definitions of variables and accounting practices, (vi) conversion into revenue allowances, (vii) robustness of the results. He concludes that benchmarking techniques are not robust and cannot completely replace detailed investigation of allowed costs. Poudineh and Jamasb (2015) point out that benchmarking results depend on the cyclicality of capital investments. Network investments are 'lumpy', usually occurring every 5-6 years, implying increased uncertainty in benchmarking analysis, because investments are mostly irreversible. As benchmarking is based on subjective and ad hoc choice it introduces regulatory risk for the companies, thereby potentially discouraging investment. On the other hand, the close relationship between the regulator and the utilities creates the potential for regulatory capture, or "gaming". Jamasb et al. (2004) found that utilities, constrained by the regulatory regime, may attempt to influence, or "game" with regards to: (i) the use of benchmarking in incentive regulation, (ii) the choice of method, model, and variables (and their weighting), (iii) the definition of variables adopted during the consultation process, and (iv) the translation of efficiency scores into X-factors.

The choice of benchmarking model serves specific regulatory objectives and furthermore affects the long-term structure of the sector, sometimes unintentionally. For example, the choice between constant or variable returns to scale models make assumptions about the potential for mergers and acquisitions. Cases of failed regulation models have been documented. In 2000-2006, Swedish DSOs experienced radical a productivity slowdown and falling profitability and efficiency, due to the failure to implement a revenue cap model (Agrell and Grifell-Tatje, 2016). Nillesen and Pollitt (2007) estimate that Dutch electricity consumers paid at least €300m for the distribution of electricity that might otherwise have been the case, due to mistakes in X-factor calculations.

2.4.2. Methods and Implications

Benchmarking techniques can be grouped in a few different manners. In CEPA (2003), benchmarking techniques are categorised into programming techniques, econometric techniques, and process (see Table 1). Alternatively, these techniques can be divided into "frontier-orientated" or "average-orientated" (Jamasb and Pollitt, 2001). The most common frontier benchmarking methods are Data Envelopment Analysis (DEA), Corrected Ordinary Least Square (COLS), and Stochastic Frontier Analysis (SFA), and the average is, for instance, the Ordinary Least Squares (OLS).

Econometric methods assume a particular functional form of the relationship between inputs and outputs as they provide parameter estimates. Econometric methods can be further categorised as deterministic or stochastic. The deterministic approaches assume that all the deviation from an estimated frontier is due to inefficiency. Under a stochastic approach, however, inefficiency is decomposed into inefficiency and measurement error.

Process techniques attempt to assess efficiency using 'bottom-up' techniques, using industry and expert experience. Under norm or average models, firms are rewarded in relation to their performance against the average cost of a group of comparable firms.

Table 1

Categorising the benchmarking techniques

Programming techniques	Linear programming approaches	Data Envelopment Analysis (DEA), Malmquist Productivity Index,
		Parametric Programming Analysis (PPA)
	Index approaches	Partial Factor Productivity (PFP)
		Total Factor Productivity (TFP)
Econometric techniques	Deterministic	Corrected Ordinary Least Squares (COLS)
	Stochastic	Stochastic Frontier Analysis (SFA)
Process approaches	Engineering economic analysis	Engineering Econometric Analysis (EEA)
	Process approaches	Process benchmarking

Source: CEPA, 2003, p.15

2.4.3. International Benchmarking

Benchmarking may be restricted by the availability of only small samples. Haney and Politt (2012) list small sample sizes as the main hindrance to performing meaningful frontier benchmarking. Adding international comparators can improve the validity of the analysis, as utilities are benchmarked against more firms. Furthermore, international comparisons enable efficiency measurement relative to the international best practice. The comparison then incorporates the frontier technical possibilities (Jamasb and Politt, 2002)

The validity of international comparisons is often undermined by the issues of the comparability of data. Different national energy offices often use various definitions of key variables, particularly costs. Furthermore, the value of capital assets in the regulatory asset base is a function of previous industry specific inflation, the taxation and capitalisation regimes, and it is therefore challenging to arrive at a number that can meaningfully be used in efficiency comparisons (Haney and Politt, 2012). Therefore, such studies are often restricted to the comparison of Opex because of the heterogeneity of capital and its availability (Jamasb and Politt, 2001). However, comparing labour costs may also require adjustments to account for pension requirements, differing social insurance costs and the degree of unionisation. The treatment of input taxes such as property taxes or public land use rights is difficult to adjust for across countries given the unclear incidence of taxation (Haney and Politt, 2012). Depending on the type of regulation and benchmarking specification, the firms will optimise by switching allowable expenditure between Capex and Opex.

Despite these issues, several notable studies have made use of international samples. Pollitt (1995) compared the effects of the public and private ownership on the performance of generation, transmission, and distribution utilities using DEA, COLS, and SFA models in a large international sample. Estache et al. (2004) used a sample of 6 Southern American countries and recommended more cross-border cooperation of regulators. Hattori et al. (2005) examined the relative performance of electricity distribution operators in Japan and the UK between 1985 and 1998. They found that the productivity gain in UK electricity distribution has been more significant than in the Japanese sector. Cullmann and von Hirschhausen (2008a) compared the efficiency of the electricity distribution operators in East European transition countries and found that privatisation had a positive effect on the technical efficiency. The Czech electricity DSOs consistently obtained the highest efficiency scores, whereas the Polish companies had the lowest efficiency scores in the region and were the most heterogeneous. Estache et al. (2008) compared the productivity development of electricity distribution systems in 12 Southern African countries between 1998 and 2005 have found comparable levels of efficiency in the region and evidence of positive effects of adopting new technologies. Countries can learn from each other as, indeed, was the case with UK being one of the forerunners in modern utility regulation, adopt new technologies and spread effective practices. International comparisons' value grows with greater cross-border collaboration and harmonisation of definitions and policies.

3. Case study: Poland and the UK

Chapter 3 presents the current regulatory frameworks in Poland and in Great Britain, in the light of the public utility literature. It compares the two countries with respect to the electricity distribution market, as well as quality of service regulation.

Poland and the United Kingdom both belong to the European Union and have not adopted the common currency. The UK, a leading trading power and financial centre, is the third largest economy in Europe after Germany and France, with GDP per capita $43,600 (PPP). Poland is the sixth-largest and one of the fastest growing economies in the EU, with $29,300 (PPP) (CIA, n.d.). There are significant differences between the levels of quality of service in Poland and the UK[9]. Since the 1989 transition, Poland made significant improvements in the efficiency of its electricity sector, but the outage duration time and frequency belong to one of the highest in Europe.

3.1. Great Britain

Electricity supply in the UK was restructured and privatised in 1990. Accompanying RPI-X regulation was introduced at privatisation, and an independent regulatory agency, Ofgem was established. Since 1995, the RPI-X price control reviews resulted in a 55% real reduction in the price of electricity distribution (Politt, 2005). In 1999, the distribution and supply activities were legally unbundled. In result, UK is one of the forerunners of the incentive regulation in the electricity sector. There are 14 licensed DSOs in Britain and each is responsible for a regional distribution services area[10]. They are owned by six different groups: Electricity North West Limited, Northern Powergrid, Scottish and Southern Energy, ScottishPower Energy Networks, UK Power Networks and Western Power Distribution. In 2000, Ofgem introduced quality of service reporting (Ofgem 2003). The current price control (referred to as RIIO-ED1) was introduced in 2015 and includes financial incentives with respect to performance in: Reliability & availability, Connections, Customer service and social obligations and Environment. Total regulated electricity distribution company revenue in 2017/2018 was around £8.8bn (Ofgem, 2017).

[9] See Figure 1 in the Appendix for the difference in the interruption times (measured by SAIDI) between Poland and the UK.
[10] See Figures 2 and 3 in the Appendix for the territorial coverage of the UK's and Poland's electricity distribution companies.

3.2. Poland

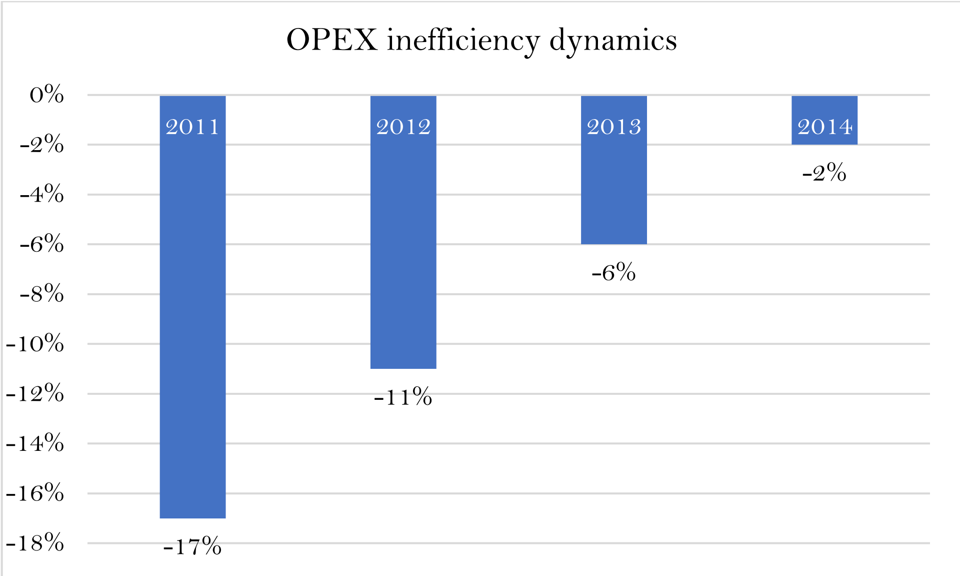
Electricity sector restructuring proved to be the most challenging part of the post-1989 economic transition process and remains mostly nationalised. Large parts of the network date back to the pre-1989 times of Polish People's Republic; Tauron Dystrybucja (2014) calculated that 25% of their network is over 40 years old. Initially, there were 33 electricity distribution companies, which were later reduced to 14, as a result of mergers. There were 172 entities registered and licensed to perform energy distribution in 2016, but there are currently five primary Distribution System Operators on the Polish market: ENEA Operator, ENERGA-Operator, PGE Dystrybucja, Innogy Stoen Operator, TAURON Dystrybucja. Four of the five legally separate DSOs are functioning within the framework of publicly traded capital groups, which are vertically integrated energy companies, where the State Treasury owns most of the packages. Only one legally unbundled DSO is owned by a company whose major shareholders are not related to the State Treasury: Innogy Stoen Operator. Distribution network charges account for one-quarter of final electricity prices in Poland[11]. Using an extensive panel dataset for 1997–2002 Cullmann and von Hirschhausen (2008b) discovered that technical efficiency increased in that period, but allocative efficiency did not. Innogy Stoen Operator, which serves Warsaw, consistently achieved the highest efficiency scores, probably as a result of servicing a high-density area. Furthermore, they demonstrated that the largest companies were more efficient and recommended more concentration. The concentration resulted in the current structure, with 5 DSOs distributing the majority of the electricity. However, as a result of the mergers, the number of companies in the sample available for benchmarking was made too low to provide sound statistical analysis in the frequentist paradigm.

The revenue of each DSO is determined by RoR regulation, which regulates the level of operating expenditures, energy losses and capital employed. URE uses a Bayesian Stochastic Frontier Analysis, developed by Makieła and Osielewski (2015), to determine the cost-effectiveness of the Polish DSOs. The study, using over forty variables for the years 2008-2014, did not find significant differences in the performance of the five DSOs, and any differences between their efficiency ratings are less than the model error. Figure 3 illustrates the average pace at which the Polish DSOs caught up with the frontier cost efficiency. However, the shortcoming of the study is that it does not incorporate measures of service quality (such as SAIDI) but focuses on minimising the losses in the system.

[11] See Figure 4 in the Appendix.

Figure 3

Cost efficiency catch-up for the 5 DSOs



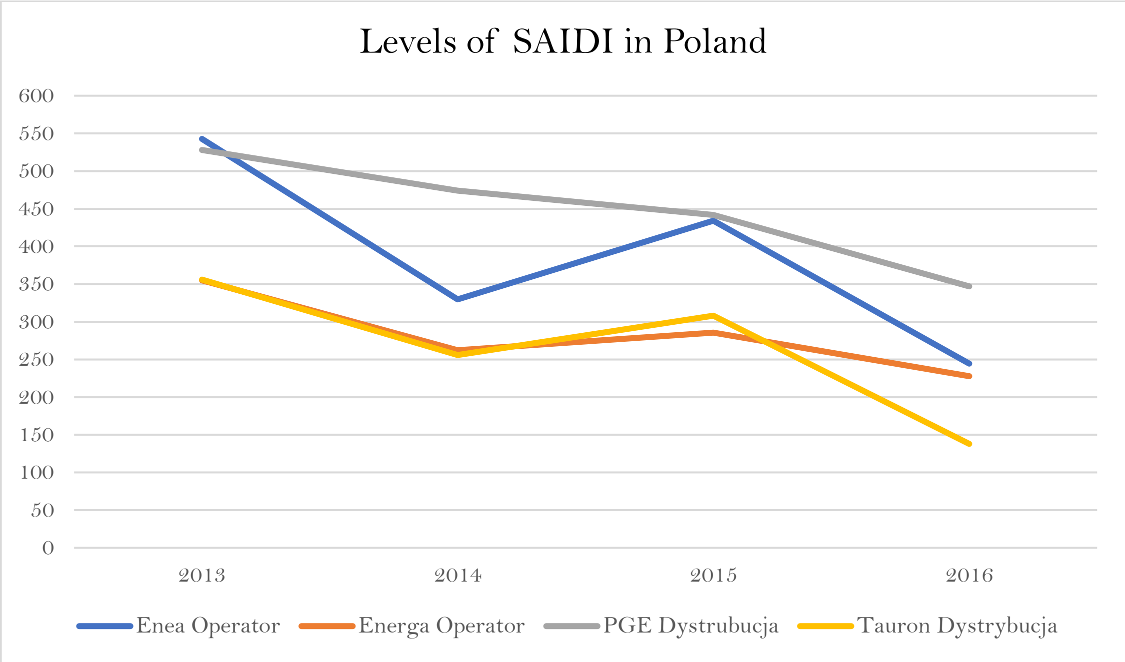
Source: Borowiak and Sroka, 2016

In 2015, the model of cost-effectiveness and technical efficiency (network losses) was updated, using, among others, benchmarking. This model was the starting point for the next regulatory period, i.e. 2016-2020 (URE, 2015c). DSOs were required to install meters in the low voltage transformer stations, so that by 2015, at least 51% of country customers were covered. In the model, annual quality improvement in terms of SAIDI, SAIFI and network connection times are specified for each DSO, with penalties and rewards for performance, incorporated in the return on the capital employed. The regulator uses historical performance to set the benchmark for each DSO. One of the aims is to achieve 50 % reduction in the SAIDI by 2020, relative to the initial 2015 level[12]. 2015 was chosen as the base year for the entire price control period. The DSOs were informed on that fact. It was, therefore, in their best interests to increase the levels of SAIDI and SAIFI, so that future years are easier targets to achieve. This is known as the "ratchet effect" described earlier. Figure 4 shows that in 2015 there was, indeed, a higher level of SAIDI for 3 out of 4 companies: Enea Operator, Energa Operator and Tauron Dystrubycja.

[12] See Figure 5 in the Appendix.

Figure 4

SAIDI for years 2013-2015 in Poland



Source: Author

The quality indicator is incorporated into the return on capital employed ($(r + \delta)$ in (1)) in a given year[13]. The total effect of the quality factor on the regulated income is to be less than 0,5 %. From the point of view of the DSO, the annual levels of SAIDI and SAIFI are crucial and do not only depend on the levels of company efforts but also on factors outside their control, such as the weather.

[13] See Equation 1 in the Appendix.

4. Methodology

This chapter describes the two models used in our study: two-stage bootstrapped Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA). In the first stage of the DEA, the bootstrapped efficiency scores are calculated and in the second they are regressed on the environmental variable COUNTRY. The SFA model is a cost function with half-normal distribution of the inefficiency term. The two models will be used to assess the performance of the companies in the sample and draw conclusions on the modelling of electricity distribution networks.

4.1. Data Envelopment Analysis

4.1.1. Traditional DEA

DEA is a non-parametric piecewise linear programming technique, which calculates the efficiency relative to a maximum likelihood estimate of an unobserved frontier, developed in Farrell (1957), Charnes et al. (1979) and Färe et al. (1985). DEA has been widely used in performance studies on hospital units, schools and engineering systems, including electricity distribution, among others[14].

Assume X, a matrix of inputs, and Y matrix of outputs. Each decision-making unit (DMU) is free to maximise its relative efficiency. The relative efficiency or the efficiency score is the ratio of the total weighed output to the total weighted input. A fully efficient firm (with a score close to unity) – or a linear combination of other firms – is the one that produces the most output using the least inputs. Thus, the efficiency frontier 'envelops' all the observations of dominated input/output combinations. The efficiency scores for the input-oriented[15] DEA model, E_i ($i \in K$ = set of observations) are found by solving the following linear program for each DMU (Bogetoft and Otto, 2010):

$$\widehat{\theta^k} = \min\{\theta \in \mathfrak{R} \big| (\theta x^k, y^k) \in \hat{T}\} \quad (k = 1, \dots, n) \tag{3}$$

Or,

[14] See Sueyoshi et al., (2017) for a recent review in energy and environment

[15] DEA models can be specified as input-oriented or output-oriented (i.e. either minimizing inputs for a given level of output, or maximising output for a given level of input). Typically, an input-oriented specification is regarded as appropriate for electricity distribution as the demand for distribution services is considered exogenous, and effectively beyond the control of distribution utilities. However, the decision to include a variable as an input or an output does not mean these are exogenous or endogenous to the firm, especially given long time frames.

$$\widehat{\theta_i} = \min_{\theta,\lambda} \theta \tag{4}$$
$$\text{s.t. } y_i + Y\lambda \geq 0,$$
$$\theta x_i - X\lambda \geq 0,$$
$$\lambda \in \wedge(\gamma).$$

where θ is a scalar and λ is a Kx1 vector of constants, which determines the T technology (i.e. γ - *returns to scale*). All models assume *free disposability* and *convexity*[16].

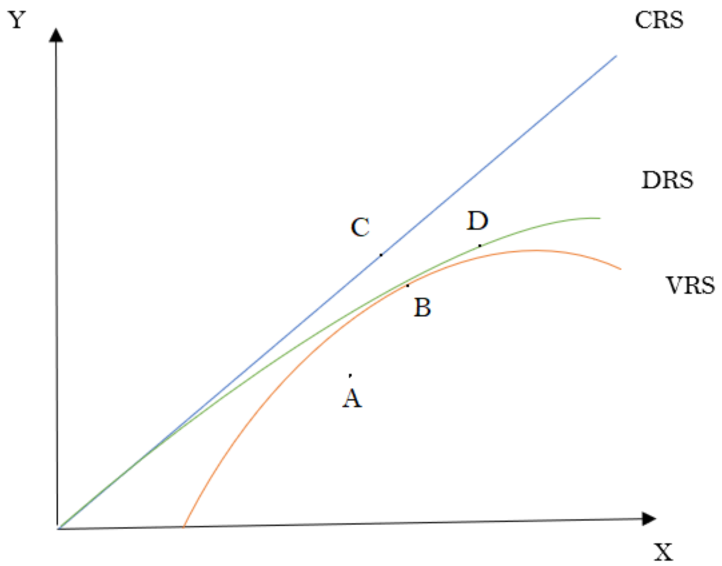
$$\wedge(\gamma)(\text{VRS}) = \{\lambda \in \mathfrak{R}_+^K \mid \sum \lambda_{k=1}^K \ \lambda^k = 1\} \tag{5}$$
$$\wedge(\gamma)(\text{DRS}) = \{\lambda \in \mathfrak{R}_+^K \mid \sum \lambda_{k=1}^K \ \lambda^k \leq 1\}$$
$$\wedge(\gamma)(\text{CRS}) = \{\lambda \in \mathfrak{R}_+^K \mid \sum \lambda_{k=1}^K \ free\} = \mathfrak{R}_+^K$$

The constant returns to scale (CRS) is the strongest assumption, where firms are assumed to be operating at their optimal size, thus assuming possible rescaling. Variable returns to scale (VRS), on the other hand, assumes no rescaling. The VRS approach compares companies with similar sizes. Some authors suggest this approach is more appropriate when the utilities are not free to choose or adapt their size (Cullman and von Hirschhausen, 2008a). DRS assumes decreasing returns to scale; it is less efficient to be large. However, it is perhaps more correct to deduce the technology using the data, rather than rely on theory and anecdotal evidence. Figure 5 shows the envelope shapes based on different technologies. Point A is inefficient under all assumptions. Point B is fully efficient under VRS, but not under DCR and CRS. Point D is fully efficient under DRS, but not under CRS. Only point C is fully efficient, based on the largest technology set CRS.

[16] See Glossary of Terms for further explanations.

Figure 5

CRS, IRS and DRS technologies



Source: Author

DEA does not require specification of production or cost function (apart from convexity assumption) or making assumptions about the distribution of the population (hence, non-parametric).

In our study, we will use a pooled cross-sectional (window) dataset. Each DMU will be assigned a number 1-4 designating the year (2013-2016). This allows for a greater sample of observations to be compared.

4.1.2. Bootstrap Data Envelopment Analysis

DEA models do not allow stochastic variations in input and output, i.e. are constructed without consideration for uncertainty, such as specification and data entry errors, or sample noise. Resultantly, it has no statistical properties and consequently leads to generate biased DEA estimates. Since the true frontier is not observed, the results are likely to be biased upwards as the estimated frontier can only be equal or worse. Bootstrapping is a computer-based statistical method, introduced by Efron (1979), where sampling is done with replacement, thus simulating the underlying data generating process (Simar and Wilson, 1998, 2000a, 2000b) and creating a bootstrap pseudodata S*.

Resampling from the observed scores, we construct an empirical sampling distribution, which tends to the true distribution. The empirical distribution is then used to estimate confidence intervals of the efficiency estimates. The resulting bootstrap distributions should be similar to the residual distribution.

A bias-corrected estimator of the true $(\tilde{\theta})$ efficiency score is:

$$\tilde{\theta} = \hat{\theta} - bias^* = 2\hat{\theta} - \bar{\theta}^* \quad (6)$$

where: $\hat{\theta}$ is the DEA estimated efficiency based on the estimated technology $\hat{T}$, $\bar{\theta}^*$ is the bootstrap estimate of the true θ, and the $bias^*$ is the bias estimated from bootstrapping. Bias correction centres the distribution of the estimator to its expected value (Tziogkidis, 2012).

Here, we present an algorithm to implement heterogeneous bootstrap DEA (Kneip et al, 2008):

1. Compute $\hat{\theta}^k$
2. Use bootstrap with smoothing from $\hat{\theta}^1, \ldots\hat{\theta}^k$ to obtain bootstrap replica $\hat{\theta}^{1*}, \ldots\hat{\theta}^{k*}$
3. Calculate bootstrapped pseudo input based on bootstrap efficiency $x^{kb} = \frac{\hat{\theta}k}{\hat{\theta}k*}x^k$.
4. Solve the DEA programme using the new set of pseudo-inputs and the same set of outputs and calculate the bootstrapped efficiency scores
5. Repeat steps 2-4 K times to calculate bootstrap estimates $\hat{\theta}^{kb}$, mean and variance of bootstrap estimates, the bias corrected estimate $\tilde{\theta}^{k*}$ and the variance.

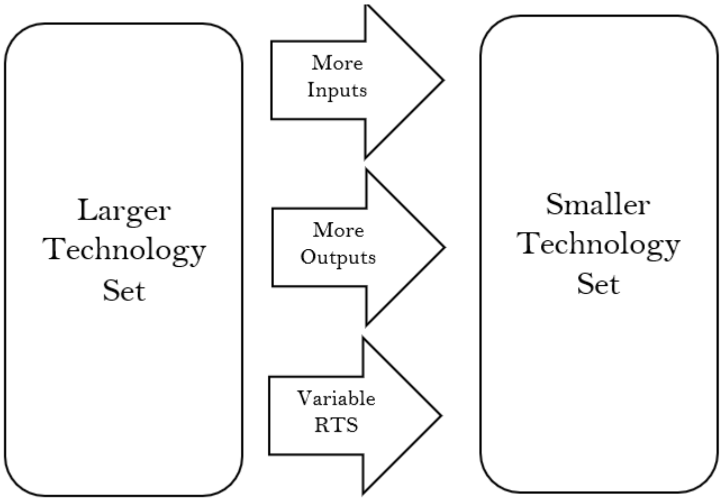
4.1.3. Hypothesis testing

One of the most important uses of bootstrapping is to test hypotheses, especially in cases where statistical inference is impossible. DEA does not provide guidance for the model specification. The choice of the best model is in the discretion of the user (Nataraja and Johnson, 2011). Smith (1997) shows that model misspecification has significant impacts on DEA efficiency estimates, especially in smaller samples.

There is a significant trade-off present in the DEA model selection. The research objective is to define a technology set T* constructed according to the minimal extrapolation principle, T* being the smallest possible subset that contains all the data and satisfies technological assumptions of convexity and free disposability (Bogetoft and Otto, 2010). In such a case, the efficiency estimates are the highest, and the ranking is supposedly the most correct.

Figure 6

Model preference



Source: Author

In general, models with more restrictions (more inputs and outputs, smaller technology set) are preferred. As such, a VRS model will be preferred to a CSR model with the same number of variables. Figure 6 presents a simple algorithm, where models with more variables dominate the ones with fewer, and where smaller technology sets dominate larger equivalents. With an increasing number of inputs and outputs, the observations in the data set are projected into a greater number of orthogonal directions and the Euclidean distance between the observations increases, which results in making the model lose its discriminatory power (Nataraja and Johnson, 2011).

Banker (1996) developed two F tests for determining the group differences, model assumptions and returns to scale. This paper will report on one of them (7). Under the null hypothesis, the two groups have the same distribution of efficiency.

$$T_{Ex} = \frac{\sum_{k \in K1} t(F^k - 1)/K_1}{\sum_{k \in K2} t(F^k - 1)/K_2} \tag{7}$$

where $t(F^k)$ is asymptotically χ^2 distributed with 2K degrees of freedom.

Simar and Wilson (2002) suggested a bootstrapping-based (SW) test, which this paper will report (8). If the H_0 is true, then S will be close to 1, and the alternative is true if it is significantly smaller than 1. The type 1 error is then Pr ($S < c_\alpha \mid H_0$)= α. If the H_0 is rejected, the model with fewer restrictions is assumed.

$$S = \frac{\sum_{k=1}^{K} E_{g1}^{k}}{\sum_{k=1}^{K} E_{g2}^{k}} \tag{8}$$

Also, a non-parametric Kolmogorov–Smirnov (KS) test, used to compare two samples of data to test the null hypothesis that the data are from the same distribution, will be reported.

In the second stage analysis, the results from the preferred model will be regressed on the environmental variable COUNTRY and DENSITY, using left-truncated regression, following Simar and Wilson (2007).

4.2. Stochastic Frontier Analysis

SFA is a parametric technique for frontier estimation, developed by Meeusen and van den Broeck (1977) and Aigner, et al. (1977). It is specified as a regression model characterised by a composite error term in which the idiosyncratic disturbance, aiming at capturing measurement error and any other noise, is included with a one-sided disturbance that represents inefficiency (Belotti et al, 2013). Following Pitt and Lee (1981), the specified panel data model is:

$$y_{it} = \alpha + x'_{it}\beta + \varepsilon_{it}, i = 1, \ldots, N, t = 1, \ldots, T \tag{9}$$

$$\varepsilon_{it} = v_{it} - u_i$$

$$v_{it} \sim N(0, \sigma_v^2)$$

$$u_i \sim N^+(0, \sigma_u^2)$$

where:

y_{it} – total cost of the i-th firm, at time t,
α – a constant,
x' – matrix of outputs of the ith firm,
β – vector of unknown parameters,
v_{it} – random (stochastic) variables assumed to be i.i.d. $v_{it} \sim N(0, \sigma_v^2)$ and independent of the u_i,
u_i – non-negative random variables assumed to account for the cost of inefficiency in production, which assumed $u_i \sim N^+(0, \sigma_u^2)$.

v_{it} is caused by stochastic disturbances, for example, unexpected expenditures for line repairs, caused by weather etc. The u_i is the degree of inefficiency or the distance from the cost frontier. Although the two components of the residual can have several different distributions, we assume normal distribution of idiosyncratic error and a half-normal distribution of inefficiency which only takes non-negative values. Therefore, the functional form is

$$Ln(TOTEX_{it}) = \alpha + \ln(x'_{it})\,\beta_1 + D_{PL}\beta_2 \qquad (10)$$

where x'_{it} is the different outputs and D_{PL} is a dummy representing COUNTRY. Furthermore, we use likelihood ratio test to compare between the different models. The resulting test statistic is distributed chi-squared, with degrees of freedom equal to the number of parameters that are constrained. The null hypothesis is that the smaller model is the "true" model, a large test statistic indicate that the null hypothesis is false.

5. Data

This chapter introduces the dataset used for the Results section (Chapter 6). The focus of this study is on the relative performance of the UK and Polish electricity distribution systems between 2013 and 2016. The dataset is panel data of 18 utilities (14 UK and 4 Polish DSOs) with 72 observations. The data for the UK were collected from the Ofgem RIIO-1 summary, whereas the data for Polish companies were taken from annual company reports and statistics. The information on one of the Polish DSOs, Innogy Stoen Operator, is entirely unavailable to public information. Therefore, we do not include it in the sample.

5.1. Choice of Variables

Whereas the modelling techniques are well established, the choice of variables of interest remains a contentious issue. The technology of electricity network service is difficult to model. Jamasb and Pollitt (2001) outlined the most widely used input and output variables. The most common inputs for electricity distribution models include the number of employees, assets (e.g. network size and transformer capacity) and expenditure (Opex, Capex and Totex). The most widely used outputs include units of energy delivered and the number of customers. Non-controllable inputs (i.e. environmental factors) include customer density and weather conditions (e.g. distance to sea, temperature, etc.), or the customer service area.

- *TOTEX* is the total expenditure sum of operating (Opex) and capital expenditures (Capex). The value given is deflated using PPI and given in US dollars PPP (base year=2010).
- *CUST* is the total number of commercial and domestic clients, given in thousands.
- *UNITS* is the total energy delivered to both commercial and domestic clients, given in GWh.
- *LENGTH* is the total length of under- and over-ground lines, given in kilometres. It can be seen as either a capital input or a proxy of the geographical extent of the service area (an output).
- *SAIDI* is the System Average Interruption Duration Index: the average outage duration per customer.
- *SAIFI* is System Average Interruption Frequency Index: the average number of interruptions per customer.
- COUNTRY is a dummy variable which takes a value of 1 when COUNTRY=Poland.

SAIFI and *SAIDI* are both undesirable outputs that violate the assumption of isotonic variables[17]. As a solution, the reciprocal of their values is taken.

The monetary value (TOTEX) is used as the sole input in all DEA models and the SFA. This approach is consistent with the current Ofgem price control. The physical variables, CUST, UNITS and LENGTH, and the quality indicators, SAIFI and SAIDI are used as outputs. The three physical variables are among the most important cost drivers of electricity distribution (Burns and Weyman-Jones, 1996; Jamasb and Pollitt, 2003).

The general rule of thumb suggests that $N \geq \max\{m^*s, 3(m+s)\}$, where N is the number of DMUs, m is the total number of inputs and s is the total number of outputs (Paradi et al. in Cooper et al., 2011). Therefore, the minimum number of observations is 20 (five outputs, one input).

The inputs and outputs seem to be strongly correlated. Table 2 shows the correlations between all variables under consideration. Predictably, SAIDI and SAIFI are strongly correlated. Also, there are strong correlations between TOTEX and LENGTH, and CUST and UNITS. Lopez (2016) shows the degree of correlation between the inputs and outputs tends to affect average efficiency scores. High correlations between inputs and outputs tend to be associated with high efficiency scores on the average. Conversely, when the correlation between inputs and outputs is close to zero, the average efficiency score of a data set is usually small. There is still no consensus in the literature as to how to overcome this problem.

Table 2

Variables correlation (Pearson correlation)

	UNIT	CUST	SAIFI	SAIDI	TOTEX	LENGTH
UNIT	1.0000000	0.9004527	0.3279096	0.2909175	0.6228943	0.6407295
CUST		1.0000000	0.6381182	0.6116550	0.8477597	0.8831486
SAIFI			1.0000000	0.9720402	0.8623000	0.8672378
SAIDI				1.0000000	0.8309880	0.8348476
TOTEX					1.0000000	0.9653197
LENGTH						1.0000000

Source: Author

[17] Traditional DEA models assume inputs and outputs to be isotonic meaning that higher inputs reduce efficiency and higher outputs increase efficiency (i.e. no undesirable outputs)

5.2. Descriptive Statistics

Table 3 presents descriptive statistics. The Polish companies tend to be larger, serving more customers and distributing more energy. This finding is predictable given the number of DSOs in both countries and the size of the population. The Polish market seems to be more highly concentrated. Quality indicators are significantly higher in Poland, by almost a factor of 8; the UK is the forerunner in quality-incorporated incentive schemes. The total expenditures in Poland are expectedly higher, as they service larger areas and customer bases. There is some indication of worse cost efficiency: using TOTEX/CUST ratio Poland still spends three times as much per customer. Also, there is a higher ratio of capital to the customer, measured by LENGTH/CUST, which may indicate lower population density[18] or overinvestment in the grid.

Table 3

Descriptive Statistics

		Min	Max	Mean	Median
Poland	TOTEX ($m)	589.8	2670.3	1848.3	2126
	CUST	2438037	5417900	4018000	4100360
	UNIT (GWh)	17.27	49.68	30.16	26.75
	LENGTH (km)	121300	285701	214925	226000
	SAIFI	2.55	5.50	3.68	3.62
	SAIDI	137.93	542.60	306.75	338.07
Great Britain	TOTEX ($m)	143.5	442.8	291	277.4
	CUST	750446	3599594	2100814	2274880
	UNIT (GWh)	7.86	35.11	21.27	21.53
	LENGTH (km)	35362	97261	56541	53005
	SAIFI	0.172206	0.81	0.46	0.52
	SAIDI	19.78239	74.68	39	40.55

Source: Author

[18] Poland: 123 people/km^2, UK: 268 people/km^2 (Statista, n.d.)

6. Results

This chapter presents the results of the eight bootstrapped DEA models and the three SFA models, and the relevant tests. Firstly, the preferred DEA model will be determined using the hypothesis tests and then presented. Secondly, the eight DEA models will be evaluated with respect to consistency of results. Thirdly, the preferred SFA model will be determined and presented, and both the DEA and SFA results will be compared in relation to other studies. Lastly, the limitations of the study will be discussed.

6.1. DEA Results

6.1.1. Model determination

Table 4 presents the eight models under consideration. The base model M1 starts with only one input and one output. Successively, an output is added, interchanged, or a more restrictive returns to scale specification is imposed.

Table 4

Model specifications

Model	Technology	TOTEX	CUST	UNITS	LENGTH	SAIDI	SAIFI
M1	CRS	I	O				
M2	CRS	I	O	O			
M3	CRS	I	O	O	O		
M4	CRS	I	O	O	O		O
M5	CRS	I	O	O	O	O	O
M6	CRS	I	O	O	O	O	
M7	VRS	I	O	O	O		
M8	DRS	I	O	O	O		

Source: Author

Table 5 presents the results of the model determination. The null hypothesis is that the variable being tested does not influence the production process. If the null is rejected, the added variable is deemed significant. The successive model is then assessed in comparison to the previous one deemed significant. For example, M2 is compared to M1, M3 is compared to M2, but M5 is already compared to M3, as M4 is insignificant.

Table 5

Model determination

H_0: $T_1 = T_2$; H_1: $T_1 \neq T_2$								
Model	Banker F-test		KS test		SW test			
	Test statistic	95%	Test statistic	Test statistic	10%	5%	2.5%	Type I error
M2	1.232	1.317	0.004*	0.840*	0.946	0.932	0.918	0.00014*
M3	1.272	1.317	0.044*	0.876*	0.951	0.941	0.931	0*
M4	1.002	1.317	0.883	0.986	0.957	0.950	0.941	0.522
M5	1.023	1.317	0.707	0.969	0.950	0.940	0.932	0.337
M6	0.797	1.317	1.000	1.106	1.054	1.042	1.030	0.673
M7	1.300	1.317	0.066	0.886*	0.925	0.917	0.908	0.002*
M8	1.215	1.317	0.186	0.905*	0.940	0.931	0.922	0.004*
Note: Test statistics marked with an asterisk* are significant at 5% level								

Source: Author

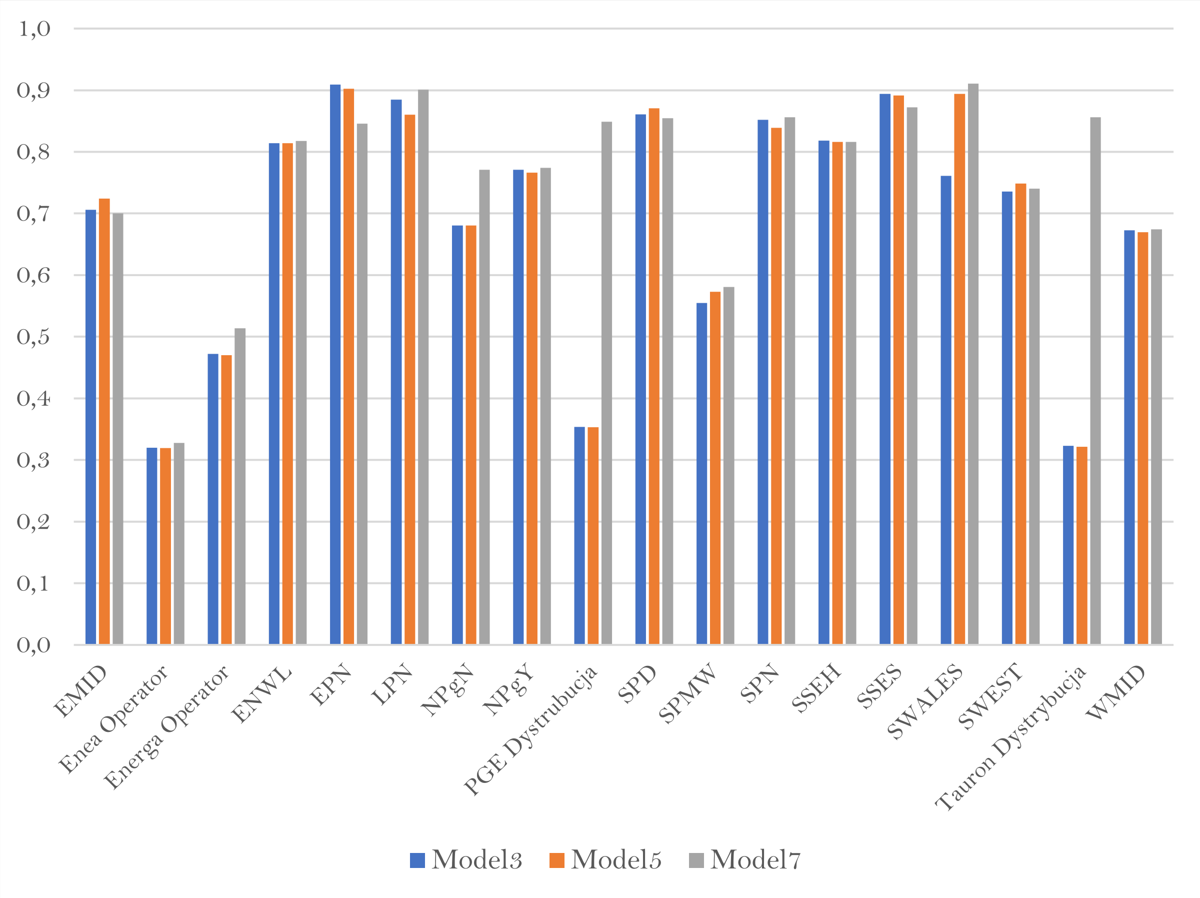
Variables CUST, UNITS and LENGTH are all significant at 5 % level in M3. The results of the tests are inconclusive, as the Banker F-test does not lead us to reject the null in cases M2-M3. However, the SW test is constructed by taking into consideration the bootstrap variations and therefore is more robust. Neither separately nor together, do SAIFI and SAIDI seem significant. Perhaps, the other output variables already explain a significant proportion of the efficiency variation.

M7 and M8 test for the returns to scale specification. Both are significant, using the SW test, but not Banker and KS tests. The VRS case, more restrictive, seems plausible given the sample data with different sized companies.

Figure 7 compares three models, M3, M5 and M7. First, two are CRS cases, and the third is the VRS version of M3. The results are consistent with the three models, supporting the conclusions from the hypothesis testing, and the claim that quality is already sufficiently explained by other variables. Variable returns to scale importantly improves the performance of two Polish companies, PGE Dystrybucja and Tauron Dystrybucja, which are the biggest gainers (50 and 53 percentage points respectively).

Figure 7

Model plot comparison (M3, M5, M7)



Source: Author

Table 6 presents the Spearman rank-order correlation coefficients for the eight models. Strong correlation (highlighted if ρ≥0.9) indicates similar results produced by the models. M5 and M4 show their strongest correlations with more parsimonious M3, and each other, supporting the earlier claim that quality indicators do not explain much of the variation. Interestingly, M6 is strongly correlated with M1 and M2, perhaps suggesting that SAIDI is somehow more linked to CUST and UNITS than SAIFI, but that would require additional testing. Predictably, the more restrictive returns-to-scale cases, M7 and M8, produce similar results. On the whole, the results of the models seem well correlated, which is a positive sign; a consistent rank scores of different models are important as they give clear indication of which utilities are more efficient, which is the main concern for the regulators (Bauer et al, 1998).

Table 6

Spearman correlation of bias-corrected efficiencies across the eight models

	M1	M2	M3	M4	M5	M6	M7	M8
M1	1.0000	0.9779	0.8427	0.8263	0.7877	0.9287	0.5191	0.5447
M2		1.0000	0.8642	0.8474	0.8018	0.9414	0.5242	0.5630
M3			1.0000	0.9889	0.9470	0.8307	0.6731	0.7177
M4				1.0000	0.9706	0.8411	0.6855	0.6989
M5					1.0000	0.8677	0.7029	0.6518
M6						1.0000	0.5958	0.5219
M7							1.0000	0.9143
M8								1.0000

Source: Author

Table 7 provides the overview of the results of the eight models. Predictably, the VRS and DRS models yield the highest results due to the flexibility and tight envelopment of the data. More parsimonious models seem to have a wider spread and generally lower average results. Including SAIDI alone (M6) seems to decrease the overall efficiency[19]. Consistently, SPD1, SPN3 and SSES1 are marked among the top results. Conversely, Energa Operator3 and Enea Operator4 are among the worst. The overall best result was achieved by LPN1 in M1 (98%). This results from the highest input/output ratio in M1.

Table 7

Descriptive statistics for the bias-corrected efficiencies (M1-M8)

Model	Min	Max	Mean	Median
M1	0.095196	0.97502858	0.440010828	0.5591
M2	0.151627	0.9674038	0.543041797	0.673857
M3	0.270873	0.9671742	0.653263922	0.765697
M4	0.269922	0.9635395	0.654314947	0.766646
M5	0.269854	0.960153	0.659367437	0.777626
M6	0.151298	0.9462297	0.546261453	0.684227
M7	0.27955	0.9548177	0.739531525	0.819772
M8	0.277162	0.9637759	0.723451315	0.806768

Source: Author

Based on the hypothesis testing, "law of parsimony" and the correlations of different models, especially the quality-incorporated M4 and M5, M3 is chosen as the preferred specification. The second model that we could prefer is M7. The decision involves a trade-off between choosing tight envelopment of data on one side and stronger correlation with quality

[19] See Table 1 in the Appendix for the full results.

incorporated models, on the other. Our choice supports a model that carries more information, with fewer variables. It includes all significant explanatory variables and is strongly correlated with M4 and M5, which include quality indicators, thus explaining their levels implicitly. It assumes constant returns to scale, which assumes that all operators are operating at their optimal levels. This is most probably correct in reality[20], as the DSO under examination have been operating for some time and have significantly developed networks. Their size, measured by LENGTH and CUST, does not change significantly year on year.

Table 8

M3 results

Unit	Rank	Bias-corrected efficiency	Bias	Efficiency	2.5%	97.5%	4-year average ($\widehat{\boldsymbol{\theta}}$)
EMID1	39	0.758	0.023	0.781	0.728	0.776	
EMID2	42	0.705	0.021	0.726	0.678	0.723	
EMID3	47	0.650	0.019	0.670	0.625	0.667	
EMID4	41	0.713	0.024	0.737	0.686	0.732	0.706
Enea Operator1	68	0.317	0.009	0.326	0.299	0.325	
Enea Operator2	63	0.345	0.011	0.356	0.326	0.355	
Enea Operator3	67	0.325	0.010	0.336	0.307	0.334	
Enea Operator4	71	0.296	0.008	0.304	0.281	0.303	0.320
Energa Operator1	58	0.489	0.024	0.513	0.457	0.511	
Energa Operator2	10	0.897	0.041	0.938	0.839	0.934	
Energa Operator3	72	0.271	0.012	0.283	0.254	0.282	
Energa Operator4	59	0.418	0.018	0.436	0.391	0.435	0.472
ENWL1	28	0.818	0.029	0.847	0.780	0.841	
ENWL2	31	0.788	0.029	0.817	0.753	0.811	
ENWL3	38	0.760	0.027	0.786	0.725	0.781	
ENWL4	11	0.895	0.038	0.934	0.853	0.926	0.814
EPN1	2	0.965	0.035	1.000	0.926	0.990	
EPN2	20	0.843	0.028	0.871	0.809	0.866	
EPN3	15	0.868	0.029	0.897	0.833	0.891	
EPN4	1	0.967	0.033	1.000	0.928	0.992	0.909
LPN1	8	0.903	0.097	1.000	0.825	0.989	
LPN2	18	0.849	0.084	0.933	0.776	0.928	
LPN3	12	0.886	0.090	0.976	0.809	0.969	
LPN4	9	0.901	0.078	0.978	0.829	0.971	0.885
NPgN1	40	0.752	0.025	0.777	0.719	0.772	
NPgN2	46	0.653	0.023	0.676	0.625	0.672	
NPgN3	52	0.628	0.022	0.650	0.600	0.645	
NPgN4	43	0.694	0.024	0.718	0.664	0.714	0.680
NPgY1	29	0.806	0.032	0.838	0.769	0.831	
NPgY2	34	0.781	0.031	0.812	0.744	0.808	

[10] In the model, however, the interpretation of returns to scale is different; VRS means that only firms of similar size are compared with each other. The two Polish companies are probably affected by this condition, being the largest in the sample.

NPgY3	45	0.676	0.027	0.703	0.644	0.699	
NPgY4	23	0.830	0.035	0.865	0.790	0.859	0.771
PGE Dystrubucja1	62	0.354	0.011	0.365	0.334	0.364	
PGE Dystrubucja2	66	0.329	0.010	0.339	0.310	0.338	
PGE Dystrubucja3	60	0.371	0.011	0.383	0.350	0.382	
PGE Dystrubucja4	61	0.363	0.011	0.374	0.343	0.373	0.354
SPD1	3	0.963	0.037	1.000	0.926	0.993	
SPD2	19	0.843	0.031	0.875	0.811	0.870	
SPD3	25	0.824	0.031	0.856	0.792	0.850	
SPD4	26	0.821	0.024	0.845	0.794	0.841	0.861
SPMW1	51	0.633	0.020	0.653	0.609	0.651	
SPMW2	55	0.558	0.018	0.575	0.536	0.573	
SPMW3	57	0.502	0.016	0.518	0.483	0.516	
SPMW4	56	0.533	0.019	0.553	0.514	0.548	0.555
SPN1	24	0.826	0.043	0.869	0.786	0.863	
SPN2	35	0.779	0.041	0.820	0.742	0.816	
SPN3	5	0.950	0.050	1.000	0.904	0.993	
SPN4	17	0.863	0.044	0.907	0.820	0.903	0.852
SSEH1	13	0.879	0.066	0.945	0.828	0.940	
SSEH2	6	0.930	0.070	1.000	0.875	0.993	
SSEH3	14	0.876	0.066	0.942	0.824	0.936	
SSEH4	53	0.625	0.042	0.667	0.587	0.664	0.818
SSES1	4	0.959	0.038	0.997	0.916	0.992	
SSES2	16	0.864	0.035	0.899	0.824	0.894	
SSES3	7	0.920	0.037	0.957	0.878	0.951	
SSES4	21	0.838	0.028	0.866	0.803	0.860	0.894
SWALES1	32	0.784	0.025	0.810	0.754	0.806	
SWALES2	30	0.804	0.026	0.830	0.773	0.826	
SWALES3	33	0.782	0.026	0.808	0.752	0.804	
SWALES4	44	0.681	0.026	0.707	0.658	0.701	0.761
SWEST1	22	0.834	0.022	0.855	0.809	0.852	
SWEST2	37	0.763	0.020	0.783	0.741	0.780	
SWEST3	27	0.819	0.021	0.840	0.795	0.837	
SWEST4	54	0.562	0.014	0.576	0.545	0.574	0.736
Tauron Dystrybucja1	64	0.339	0.013	0.352	0.323	0.350	
Tauron Dystrybucja2	65	0.333	0.012	0.345	0.317	0.344	
Tauron Dystrybucja3	70	0.307	0.012	0.318	0.292	0.317	
Tauron Dystrybucja4	69	0.315	0.013	0.328	0.300	0.326	0.323
WMID1	36	0.768	0.025	0.793	0.735	0.788	
WMID2	49	0.644	0.021	0.665	0.618	0.661	
WMID3	48	0.650	0.021	0.671	0.623	0.666	
WMID4	50	0.636	0.022	0.658	0.608	0.652	0.672

Source: Author

Table 8 presents the results of the preferred model (M3). DMUs tend to achieve consistent results over the four-year period. An exception is Energa Operator 2, which achieved almost 90% efficiency and later dropped to less than 30%, achieving the worst rank in the sample.

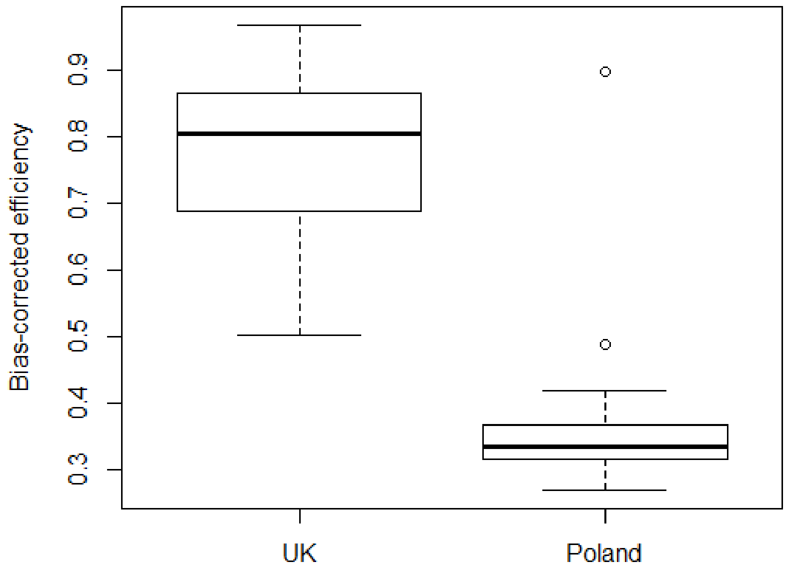
The overall best performer is EPN4 with 97% efficiency. There is no clear time trend of productivity development, as expected in such a short panel[21].

6.1.2. Second-stage testing

In the following analysis we address a structural variable: COUNTRY. It is reasonable to assume that the two countries have significant structural differences, available technologies and different legacies of the past, which are currently outside their decision framework. There is a noticeable difference in the performance between the observations of the Polish and British utilities, with the latter achieving higher results. Figure 8 shows the boxplots of the two subsamples. The performance of Energa Operator 2, previously identified as the largest one-time gainer, is a clear outlier, potentially signalling a data entry mistake.

Figure 8

Subsamples comparison; Boxplot of bias-corrected efficiencies for Poland and the UK



Source: Author

To test whether there is in fact evidence to conclude that the two subsamples are statistically different we employ Welch t-test and a truncated regression (Tables 9 and 10) on the bias-corrected efficiencies.

The t-test verifies whether the two samples have unequal means and the regression plots the utilities' efficiencies on the country of origin. In both cases, we conclude that the relationship is significant. We reject the null in the t-test and conclude the mean of the Polish subsample

[21] See Figure 6 in the Appendix for the productivity development over time.

is 52% of the British. A similar conclusion flows from the regression, where the coefficient is negative and significant, indicating that on average Polish companies are 42 % points less efficient.

Table 9

Welch t-test for unequal means

H_0: Mean in group 0 = Mean in group 1; H_1: Mean in group 0 ≠ Mean in group 1				
	Statistic	2.5%	97.5%	p-value
Welch t-test	t = 11.393	0.3324	0.4735	< 2.2e-16
Mean in group 0	0.7822584	Mean in group 1	0.3793009	
Note: Group 0 = UK, group 1 = Poland				

Source: Author

Table 10

Left-Truncated Regression (Bias-corrected efficiencies on COUNTRY)

Left-Truncated Regression			
	Coefficient	Std. Error	t-value
(Intercept)	0.803117	0.023197	34.6213*
COUNTRY	-0.42382	0.041289	-10.2646*
sigma	0.136627	0.014853	9.1989*
marked with an asterisk (*) are significant at 5% level.			

Source: Author

6.2. SFA Results

In this section, we present the results of the SFA models. First, we compared translog and Cobb-Douglas formulation. There was significant evidence to reject the more complicated model in favour of the parsimonious Cobb-Douglas formulation[22].

Table 11 reports on the results of the unconstrained SFA model. The intercept, SAIDI, SAIFI, UNITS and CUST, are strongly insignificant. SAIFI has an unexpected sign: it suggests that decreasing the level of SAIFI decreases the cost.

[22] See Table 2 in the Appendix for the LR.

Table 11

Results of the unconstrained SFA Model 1

ln(TOTEX)		Coef.	Std. Err.	z	P>|z|	95% C.I.	
Frontier β	ln(UNITS)	0.226	0.306	0.740	0.460	-0.374	0.826
	ln(CUST)	0.312	0.321	0.970	0.331	-0.318	0.942
	ln(SAIDI)	0.043	0.111	0.380	0.700	-0.174	0.260
	ln(SAIFI)	-0.058	0.154	-0.370	0.708	-0.359	0.244
	ln(LENGTH)	0.295	0.177	0.095	0.095	-0.052	0.641
	(Intercept)	-2.902	3.586	-0.510	0.610	-9.821	4.017
	D_{PL}(COUNTRY)	1.071	0.200	5.350	0.000*	0.679	1.463
δ^2	Var. of u	0.027	0.016	1.650	0.098	-0.005	0.059
	Var. of v	0.022	0.004	5.130	0.000	0.013	0.030

Source: Author

Table 12 presents the results of the constrained Model 3 (excluding quality indicators and UNITS). A constrained Model 2 (excluding quality indicators) is not reported[23]. The likelihood ratio test leads us to prefer Model 3. In Model 2, ln(UNITS) and ln(CUST) are insignificant. This is likely to result from the high correlation between the two variables (see Table 2). In Model 3[24], all variables are significant at the 10 % level. The LR test in Table 13 suggests that dropping ln(UNITS) did not affect the goodness-of-fit. The coefficient of D_{PL}(COUNTRY) is significant and can be interpreted as follows: the value of TOTEX will increase by 180% when D = 1 and decrease by 380% ((exp(±β) - 1) * 100) when D = 0, suggesting that Poland is around 4 times less efficient. Other coefficients can easily be interpreted: if we change X' by one percent, we would expect Y to change by β percent. The coefficient of ln(CUST) is the largest of outputs, at 0,529. The test for CRS is failed at 5%, meaning we should not constrain ln(LENGTH) + ln(CUST) = 1, and that there is evidence of variable returns to scale, approximately equal 0.82 (0.529 + 0.292).

[23] See Table 3 in the Appendix for the Model 2 results.

[24] See Table 4 in the Appendix for the efficiency scores in Model 3.

Table 12

Results from the constrained SFA Model 3 (excluding UNITS, SAIFI and SAIDI)

ln(TOTEX)		Coef.	Std. Err.	z	P>|z|	95% C.I.	
Frontier β	ln(CUST)	0.529	0.114	4.64	0.000*	0.305	0.752
	ln(LENGTH)	0.292	0.150	1.94	0.052	-0.003	0.587
	(Intercept)	-5.338	1.166	-4.57	0.000*	-7.624	-3.051
	D_{PL}(COUNTRY)	1.028	0.167	6.18	0.000*	0.702	1.354
δ^2	Var. of u	0.030	0.017	1.74	0.082	-0.004	0.065
	Var. of v	0.021	0.004	5.16	0.000*	0.013	0.029

Source: Author

Table 13

Likelihood ratio tests

Likelihood ratio (LR) tests		
Test hypothesis	LR chi2	Prob > chi2
Model 2 = Model 3	(2) 0.40	0.5255
Model 3: Constant Returns to Scale	(1) 3.11	0.0779
Note: The value given in (brackets) denotes the degrees of freedom in the LR test		

Source: Author

6.3. Comparison of Results

In this section, we compare the results of the two methodologies and use external results as a means of validation.

The two methodologies produce similar results and conclusions. The preferred SFA and DEA models both include CUST and LENGTH, and conclude that COUNTRY is significant in explaining the cross-country differences. They are also in agreement with respect to the insignificance of both quality indicators, SAIDI and SAIFI. It may be argued that the quality indicators are reflective of the current state of the network, which is already reflected by other variables, such as LENGTH.

Makieła (2018) used the same dataset and Bayesian SFA[25] with a number of different error distribution assumptions and specifications. The results of the preferred models can be found in the Appendix. He finds that a normal-half-normal time-invariant model with the same

[25] Using the methodology described in Makieła, K., Osiewalski, J. (forthcoming). See Tables 5 and 6 in the Appendix.

explanatory variables (CUST, LENGTH, COUNTRY) performs the best based on the CAME estimator. Moreover, he finds that the translog formulation leads to a det(A) ~ 0, therefore impossible to calculate. Using a Varying Effects Model[26], he also includes a dummy variable to account for the changes in the regulatory framework in Poland in 2015-2016, which is significant. It leads us to conclude that Polish companies did react to the imposition of the new system by reducing their quality indicators, which "cost" them a 143% increase in the TOTEX.

Table 14 presents the correlation matrix of results from the preferred DEA model (M3), SFA model (Model 3) and Bayesian SFA. The results are well correlated. The coefficients for ln(LENGTH) and ln(CUST) are similar. Coefficient of D_{PL}(COUNTRY) is likely to be affected by the introduction of additional variable D(QUALITY_PL). Unsurprisingly, the two SFA models are the most strongly correlated. The VED model is very strongly correlated with DEA; the reasons for which require further investigation.

Table 14

Spearman rank correlation matrix of efficiencies calculated using the two models (DEA and SFA) and two external (Bayesian SFA and VED)

	SFA	DEA	Bayesian SFA	VED
SFA	1.0000			
DEA	0.5651	1.0000		
Bayesian SFA	0.7220	0.6950	1.0000	
VED	0.4871	0.9617	0.6067	1.0000

Source: Adopted from Makieła (2018)

6.4. Limitations

There are significant issues with the presented methodology. Firstly, there is the problem of international comparisons, explained in Chapter 2. On one side, a more extensive dataset tends to produce results that are consistent with the underlying data generating process, but with large samples the homogeneity of the dataset may decrease and external factors, beyond the scope of the study, enter into consideration. For example, the dummy variable COUNTRY captures a number of different effects, one of which are the different types of ownership between Polish and British companies. Because we exclude Innogy Stoen, we cannot investigate the effects of different types of ownership on the performance of the Polish companies.

[26] Based on Koop et al. (1994, 1997)

Secondly, there might be an issue of missing variables. To perform a reliable efficiency, exercise a greater number of explanatory variables should be considered, for example, the number of transformers, meters or energy delivered to different customer groups. Furthermore, Totex-only approach contains fewer cost drivers than the disaggregated approaches, leading to a less intuitive relationship between cost drivers and costs, and requires some consideration of the inherited characteristics of the network and the previous spend. However, such information is not publicly available. Furthermore, the data on Polish companies often rely on imprecise information, for example in case of the length of the lines. Such errors-in-variables decreases the explanatory power of the models and affects the results, especially in the DEA.

There are other more robust approaches available for determining the validity of individual variables in the DEA, such as the Efficiency Contribution Measure, which may produce more robust results, particularly in larger data sets. Other specifications of the DEA, including statistical noise, for example, are also available but have not yet been adequately understood and the software is not available.

The SFA suffers from a relatively small sample to perform meaningful statistical modelling. In the light of that, the only viable solution is the Bayesian SFA, which allows for small-sample comparisons. However, given the results from Makieła (2018), we conclude that the method chosen is correct and the convergence of the results produced by both methodologies constitutes evidence for that.

7. Conclusions

This study compares the performance of the 18 British and Polish Distribution System Operators and the corresponding regulatory frameworks. Our analysis makes a number of contributions to the existing literature. Firstly, it successfully applies two-stage semi-parametric bootstrapped DEA in the international context of electricity distribution. Secondly, it quantifies the existing performance gap between the Polish and British companies and finds it significant. Thirdly, it compares the results with two classes of SFA models and finds consistent results with regards to model specification. Lastly, it describes the regulatory regimes in Poland and Great Britain, bringing together the two frameworks and highlighting the areas that require further examination.

Two-stage semi-parametric bootstrapped DEA allows for statistically sound hypothesis testing. Testing for group means and regressing on the environmental variables has been done in the past, but to our knowledge no such study exists in the electricity distribution with an international sample. Furthermore, the study confirms those conclusions with SFA and Bayesian SFA, which allows for the most robust small sample comparisons. There is still no consensus as to the precise methodology of model determination in the different specifications of the DEA. Our results are well correlated and converge on the choice of explanatory variables, but not on the returns to scale specification. In summary, the quality indicators are not significant in determining the level of TOTEX, which is in accordance with previous studies by Makieła and Osiewalski (2016). Our finding contradicts the findings of Giannakis et al. (2003) in that it finds that the companies that are cost efficient may also be the most efficient with regards to quality of service. Econometric analysis allows us to discern the variables that account for the majority of the variations in the efficiency scores. The conclusions of this paper are in accordance with Burns and Wayman-Jones (1996) and Makieła and Osiewalski (forthcoming) in favouring parsimonious models with relatively few explanatory variables. In practice, that does not mean that the regulator should not set targets for costs, as well as quality of service. Taking the viewpoint of the consumers, households and businesses, the levels of SAIFI and SAIDI are of crucial importance and therefore deserve separate treatment. More research could be performed on customers' willingness-to-pay, to calculate the actual desired levels of quality in the Polish market. This is confirmed by empirical evidence from other studies, such as Ter-Martirosyan (2003) and Giannakis et al. (2005) cited earlier. However, it may suggest that benchmarking of utilities is relatively simpler in later stages of the regulatory process, as all variables are strongly correlated.

International comparisons can never be implemented within a regulatory framework, unless there was a high degree of harmonisation between the countries in the sample. Nevertheless, it is useful to establish the differences and similarities for the purposes of future progress, and to understand the performance gaps that exist between countries. We find that on average Polish companies are 50 % less efficient using the DEA model. In SFA, the dummy variable COUNTRY was significant and increased the levels of TOTEX by 180% when D=1. This implies significant performance deficiencies on the Polish side. However, from a practical perspective these findings present little value, as they do not give any indication of what could be done to improve the current status quo. Benchmarking provides a contextual framework for analysing the difference in performance, but implementation requires practical policymaking and engineering expertise. The values obtained from the exercise are descriptive and do not provide any exact real measure that could be implemented. However, the importance of this study lies in contextualising the performance of the DSOs in the international sample. By allowing for an international comparison we introduce the global best practice into the sample.

To improve future international comparison, it is necessary to allow for the comparability of data between countries by harmonising definitions of variables. Data collection and transparency standards could be imposed by the Polish regulator, similar to that in Britain. For example, an open source database with all the monitored variables could be made available for the public. That would allow for more research to take place and create external pressure for efficiency. This study pointed a number of research areas that need further examination. For example, Figure 5 in Chapter 3 highlighted a potential "gaming", or the "ratchet effect", with respect to SAIFI and SAIDI. More data, including personal accounts and weather variables, would be required to verify whether it is a convincing claim, and if really is there evidence of the "ratchet effect". Moreover, it would be interesting to investigate the effects of the new incentive regulation on the levels of investment in the grid.

Collaboration between regulators is already taking place on an international level, which helps diffuse effective practices. However, each regulator faces their own individual set of conditions. The theory finds that no "size fits all", and therefore the practices lend themselves to ambiguity in implementation. Electricity distribution companies have enjoyed a "quiet life" of natural monopolies. However, technological change puts pressure on the utilities to invest in frontier solutions, whilst maintaining both cost and quality of service efficiency. It provides a challenge for both the companies and the regulators that can only be met in an environment that is open to public scrutiny and dissemination of new knowledge.

Note: Bootstrapped Data Envelopment Analysis was performed using R package "Benchmarking" (Bogetoft and Otto, 2010) and the Stochastic Frontier Analysis was performed using "sfpanel" (Belotti et al, 2013) in STATA.

References

Aigner, D., Lovell, C. K., and Schmidt, P. (1977) 'Formulation and estimation of stochastic frontier production function models', *Journal of econometrics*, *6*(1), 21-37.

Ajodhia, V. and Hakvoort, R. (2005) 'Economic regulation of quality in electricity distribution networks', *Utilities Policy*, 13(3), pp.211-221.

Alexander, I., and Irwin, T. (1997) 'Price caps, rate-of-return regulation, and the cost of capital', Note No. 87. World Bank, Washington, DC

Averch, H. and Johnson, L. (1962) 'Behavior of the Firm under Regulatory Constraint', *American Economic Review*, 52: 1059-69.

Balazs E., (2009) 'Infrastructure Investment in Network Industries: The Role of Incentive Regulation and Regulatory Independence', *CESifo Working Paper Series 2642*, CESifo Group Munich.

Banker, R. D. (1996). 'Hypothesis tests using data envelopment analysis', *Journal of productivity analysis*, 7(2-3), 139-159.

Bauer, P., Berger, A., Ferrier, G., and Humphrey, D. B. (1998) 'Consistency conditions for regulatory analysis of financial institutions: a comparison of frontier efficiency methods', *Journal of Economics and business*, 50(2), 85-114.

Belotti, F., Daidone, S., Ilardi, G., and Atella, V. (2012) 'Stochastic Frontier Analysis Using Stata', *CEIS Working Paper No. 251.*

Besanko, D., and Spulber, D. F. (1992) 'Sequential-equilibrium investment by regulated firms', *The Rand Journal of Economics*, 153-170.

Bogetoft, P., and Otto, L. (2010) *Benchmarking with DEA, SFA, and R*. Springer Science & Business Media.

Borowiak, W., and Sroka M. (2016) "Strategia regulacji jakościowej OSD. Należy rozmawiać", Departament Usług Operatorskich i Taryf Enea Operator Sp. z o.o. Energia Elektryczna. 5/2016. http://www.cire.pl/pliki/2/nalezyrozmawiac_052016.pdf

Burns, P., and Weyman-Jones, T. G. (1996) 'Cost functions and cost efficiency in electricity distribution: a stochastic frontier approach', *Bulletin of Economic Research*, 48(1), 41-64.

Cambini, C., and Rondi, L. (2010) 'Incentive regulation and investment: evidence from European energy utilities', *Journal of Regulatory Economics*, 38(1), 1-26.

CEER (2015) 'CEER Benchmarking Report 5.2 on the Continuity of Electricity. Supply Data', Council of European Energy Regulators. https://www.ceer.eu/documents/104400/-/-/cbc48e6a-5d5e-a170-ae1d-7b7b298d46a4 (Accessed 2 Mar. 2018).

Cena prądu (n.d.) "Cena prądu w Polsce", http://www.cenapradu.strefa.pl/

CEPA (2003) 'Background study on the use of benchmarking to assess efficiency for the 2005 Distribution Price Control Review', Cambridge, Ofgem

CEER (2016) '6th CEER Benchmarking Report on the quality of electricity and gas supply', Council of the European Regulators, Brussels

Charnes, A., Cooper, W. W., and Rhodes, E. (1978) 'Measuring the efficiency of decision making units', *European journal of operational research*, 2(6), 429-444.

Charnes, A., Cooper, W. W., and Rhodes, E. (1979). 'Measuring the efficiency of decision-making units', *European journal of operational resear*ch, 3(4), 339.

CIA (n.d.) "The World Factbook", Central Intelligence Agency: https://www.cia.gov/library/publications/the-world-factbook/geos/pl.html

Cooper, W. W., Seiford, L. M., and Zhu, J. (2011) *Handbook on data envelopment analysis* (Vol. 164). Springer Science & Business Media.

Cullmann, A., and von Hirschhausen, C. (2008a) 'Efficiency analysis of East European electricity distribution in transition: legacy of the past?', *Journal of Productivity Analysis*, 29(2), 155.

Cullmann, A., and Von Hirschhausen, C. (2008b) 'From transition to competition: Dynamic efficiency analysis of Polish electricity distribution companies', *Economics of Transition*, 16(2), 335-357.

Efron, (1979) 'Bootstrap Methods: Another Look at the Jackknife' *Annals of Statistics* 7, no. 1, 1--26.

Energy Network Association (n.d.) "Electricity Distribution Map" http://www.energynetworks.org/info/faqs/electricity-distribution-map.html (accessed 1st March 2018)

Estache, A., Tovar, B., and Trujillo, L. (2008) 'How efficient are African electricity companies? Evidence from the Southern African countries'. *Energy policy*, 36(6), 1969-1979.

Färe, R., and Grosskopf, S. (1985) 'A nonparametric cost approach to scale efficiency'. *The Scandinavian Journal of Economics*, 594-604.

Farrell, M. J. (1957) 'The measurement of productive efficiency', *Journal of the Royal Statistical Society*. Series A (General), 120(3), 253-290.

Farsi, M. and Filippini, M. (2004) 'Regulation and measuring cost efficiency with panel data models: application to electricity distribution utilities', *Review of Industrial Organization*, 25, 1–19.

Farsi, M., Filippini, M., & Greene, W. (2006) 'Application of panel data models in benchmarking analysis of the electricity distribution sector', *Annals of Public and cooperative Economics*, 77(3), 271-290.

Filippini, M. (1998) 'Are municipal electricity distribution utilities natural monopolies?', *Annals of public and cooperative economics*, 69(2), 157-174.

Freixas, X., Guesnerie, R., and Tirole, J. (1985) 'Planning under incomplete information and the ratchet effect', *The Review of Economic Studies*, 52(2), 173-191.

Gal-Or, E. and Spiro, M. (1992) 'Regulatory regimes in the electric power industry: Implications for capacity', *Journal of Regulatory Economics*, 4(3), pp.263-278.

Giannakis, D., Jamasb, T. and Pollitt, M. (2005) 'Benchmarking and Incentive Regulation of Quality of Service: An Application to the UK Electricity Distribution Networks', *Energy Policy* 33(17): 2256-2271

Gilsdorf, K., (1995) 'Testing for Subadditivity of Vertically Integrated Electric Utilities', *Southern Economics Journal* 62 (14)

Salvanes, K., and Tjøtta S. (1998) 'A test for natural monopoly with application to Norwegian electricity distribution.' *Review of Industrial Organization* 13.6: 669-685.

Haney, A., and Pollitt, M. (2013) 'International benchmarking of electricity transmission by regulators: A contrast between theory and practice?', *Energy Policy*, Volume 62, Pages 267-281, ISSN 0301-4215, Honkapuro, S. (2008) 'Performance benchmarking and incentive regulation–considerations of directing signals for electricity distribution companies'. Lappeenranta University of Technology.

Jamasb, T., and Pollitt, M. (2000) 'Benchmarking and regulation: international electricity experience', *Utilities policy*, 9(3), 107-130.

Jamasb, T., and Pollitt, M. (2003) 'International benchmarking and regulation: an application to European electricity distribution utilities', *Energy policy*, 31(15), 1609-1622.

Jamasb, T., and Pollitt, M. (2007) 'Incentive regulation of electricity distribution networks: Lessons of experience from Britain. *Energy Policy*, 35(12), 6163-6187.

Jamasb, T., Nillesen, P. and Pollitt, M. (2004) 'Strategic Behaviour under Regulatory Benchmarking', *Energy Economics* 26(5): 825-843.

Jasmsb, T. and M. Pollitt (2001) 'Benchmarking and Regulation: International Electricity Experience', *Utilities Policy*, 9: 107-130.

Joskow, P. L. (2008) 'Incentive regulation and its application to electricity networks', *Review of Network Economics*, 7(4).

Joskow, P.J., and Schmalensee, R., (1986) 'Incentive regulation for electric utilities', *Yale Journal on Regulation*. 4 (1), 1–49.

Kahn, A.E. (1971) *The Economics of Regulation: Principles and Institutions*, Wiley: NewYork.

Kneip, A., Simar, L., and Wilson, P. W. (2008) 'Asymptotics and consistent bootstraps for DEA estimators in nonparametric frontier models', *Econometric Theory*, 24(6), 1663-1697.

Koop, G., Osiewalski J. and Steel, M. (1994) 'Hospital efficiency analysis through individual effects: A Bayesian approach', *CentER Discussion Paper* 9447.

Koop, G., Osiewalski J. and Steel, M. (1997) 'Bayesian efficiency analysis through individual effects: Hospital cost frontiers', *Journal of Econometrics* 76(1–2): 77–105.

Kumbhakar, S. C., and Hjalmarsson, L. (1998) 'Relative performance of public and private ownership under yardstick competition: electricity retail distribution', *European Economic Review*, 42(1), 97-122.Pages 97-122,

Kwoka, J.E. (2005) 'The Comparative Advantage of Public Ownership: Evidence from U.S. Electric Utilities', *Canadian Journal of Economics*, 38 (2):622-640

Liston, C. (1993). 'Price-cap versus rate-of-return regulation', *Journal of Regulatory Economic*s, 5(1), 25-48.

Lyon, T., and Mayo, J. (2005) 'Regulatory Opportunism and Investment Behavior: Evidence from the U.S. Electric Utility Industry', *The RAND Journal of Economics*, 36(3), 628-644.

Makieła, K. (2018) 'Bayesian SFA: Results for Poland and the UK', *Mimeo*. Unpublished

Makieła, K., Osiewalski, J. (forthcoming) 'Cost Efficiency Analysis of Electricity Distribution Sector under Model Uncertainty', *The Energy Journal*, *39*(4).

Meeusen, W., and van Den Broeck, J. (1977) 'Efficiency estimation from Cobb-Douglas production functions with composed error', *International economic review*, 435-444.

Nataraja, N. and Johnson, A. (2011) 'Guidelines for using variable selection techniques in data envelopment analysis', *European Journal of Operational Research*, 215(3), pp.662-669.

Nillesen, Paul and Pollitt, Michael, (2007) 'The 2001-3 electricity distribution price control review in the Netherlands: regulatory process and consumer welfare', *Journal of Regulatory Economics*, 31, issue 3, p. 261-287,

Ofgem (2009) "Regulating Energy Networks for the Future: RPI-X@20 History of Energy Network Regulation", Office of Gas and Electricity Markets https://www.ofgem.gov.uk/ofgem-publications/51984/supporting-paper-history-energy-network-regulation-finalpdf (Accessed 3 Mar. 2018).

Ofgem. (2017). "RIIO electricity distribution annual report 2015-16" Office of Gas and Electricity Markets, https://www.ofgem.gov.uk/publications-and-updates/riio-electricity-distribution-annual-report-2015-16 (Accessed 1 Mar. 2018)

Ogden, S. G. (1997). 'Accounting for organizational performance: the construction of the customer in the privatized water industry'. *Accounting, Organizations and Society*, 22(6), 529-556.

Ordyna, P. (2014) "Regulacja jakościowa z perspektywy Operatora Systemu Dystrybucyjnego" Krakow: Tauron Dystrybucja http://piknik.tauron-dystrybucja.pl/SiteCollectionDocuments/materialy-2014/P-Ordyna-Regulacja-jakosciowa-z-perspektywy-OSD.pdf

Osiewalski J., and Makieła K. (2015) 'Koncepcja ustalania wybranych elementów kształtujących Przychód Regulowany OSD, którzy dokonali z dniem 1 lipca 2007r. rozdzielenia działalności – model kosztów operacyjnych i różnicy bilansowej', Kraków 2015.

Pitt, M., and Lee, L.F. (1981) 'The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry', *Journal of Development Economics* 9(1/August), 43-64.

Pollitt, M. (2005) 'The role of efficiency estimates in regulatory price reviews: Ofgem's approach to benchmarking electricity networks', *Utilities policy*, 13(4), 279-288.

Pollitt, M.G. (1995) *Ownership and performance in electric utilities: the international evidence on privatization and efficiency*. Oxford: Oxford Institute for Energy Studies.

Poudineh, R. and Jamasb, T. (2015) 'A new perspective: investment and efficiency under incentive regulation', *Energy journal*., 36 (4). pp. 241-263.

Price, C. W., Brigham, B., and Fitzgerald, L. (2002) 'Service quality in regulated monopolies'. University of East Anglia Centre for Competition and Regulation, *CCR Working Paper, 02*-4.

PTPiREE (2017) "Energetyka przesyłowa i dystrybucyjna: Raport", Polskie Towarzystwo Przesyłu i Rozdziału Energii Elektrycznej, http://www.ptpiree.pl/news/2017-05-24/raport_ptpiree_07.07.2017.pdf

Santos, S., Amado, C. and Rosado, J. (2011) 'Formative evaluation of electricity distribution utilities using data envelopment analysis', *Journal of the Operational Research Society*, 62(7), pp.1298-1319.

Shleifer, A. (1985) 'A theory of yardstick competition', *The RAND Journal of Economics*, 319-327.

Simar, L., and Wilson, P. W. (1998) 'Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models', *Management science*, *44*(1), 49-61.

Simar, L., and Wilson, P. W. (2000a) 'A general methodology for bootstrapping in non-parametric frontier models', *Journal of Applied Statistics*, 27(6), 779-802.

Simar, L., and Wilson, P. (2000b) 'Statistical inference in nonparametric frontier models: The state of the art', *Journal of productivity analysis*, *13*(1), 49-78.

Simar, L., and Wilson, P. (2002) 'Non-parametric tests of returns to scale' *European Journal of Operational Research*, 139(1), 115-132.

Simar, L., and Wilson, P. (2007) 'Estimation and inference in two-stage, semi-parametric models of production processes', *Journal of econometrics, 136*(1), 31-64.

Smith, P. (1997) 'Model misspecification in data envelopment analysis', *Annals of Operations Research*, 73, 233-252.

Spence, A. (1975) 'Monopoly, quality, and regulation', *The Bell Journal of Economics*, 417-429.

Statista (n.d.) "Population density in the United Kingdom (UK) as of mid-2015 (people per sq. km), c. UK population density, by country 2015" | Statistic. https://www.statista.com/statistics/281322/population-density-in-the-united-kingdom-uk-by-country/

Stern, J. (2003) 'What the Littlechild Report actually said', In *PROCEEDINGS-CRI* (pp. 7-30).

Sueyoshi, T., Yuan, Y., and Goto, M. (2017) 'A literature study for DEA applied to energy and environment', *Energy Economics*, 62, 104-124.

Ter-Martirosyan, A. (2003) 'The effects of incentive regulation on quality of service in electricity markets', Dpt. Of Economics, George Washington University, *Working Paper*.

Tziogkidis, P. (2012) 'Bootstrap DEA and hypothesis testing', *Cardiff Economics Working Papers*, No. E2012/18

Union, E. (2009). Directive 2009/72/EC of the European parliament and of the council of 13 July 2009 concerning common rules for the internal market in electricity and repealing directive 2003/54/EC. Off. J. Eur. Union, 52(L211), 55-93.

URE (2015a) 'Koszty operacyjne dla Operatorów Systemów Dystrybucyjnych na lata 2016-2020', Departament Rynków Energii Elektrycznej i Ciepła, Warszawa: Urząd Regulacji Energetyki

URE (2015b) 'Regulacja Jakośćiowa w latach 2016-2020 dla Operatorów Systemów Dystrybucyjnych', Departament Rynków Energii Elektrycznej i Ciepła, Warszawa: Urząd Regulacji Energetyki

URE (2015c) 'Strategia Regulacji Operatorów Systemów Dystrybucyjnych na lata 2016-2020', Departament Rynków Energii Elektrycznej i Ciepła, Warszawa: Urząd Regulacji Energetyki

Vogelsang, I. (2002) 'Incentive Regulation and Competition in Public Utility Markets: A 20-Year Perspective', *Journal of Regulatory Economics*, 22(1), pp.5-27.

Walls, S. (2018) "Totex modelling – part 1" Ofgem. https://www.ofgem.gov.uk/ofgem-publications/47333/enwlwpdpresentationontotex-pdf (Accessed 1 Mar. 2018).

Weitzman, M. L. (1980) 'The" ratchet principle" and performance incentives', *The Bell Journal of Economics*, 302-308.

Zhang, Y., and Bartels, R. (1998) 'The effect of sample size on the mean efficiency in DEA with an application to electricity distribution in Australia, Sweden and New Zealand', *Journal of Productivity Analysis*, 9(3), 187-204.

Appendix

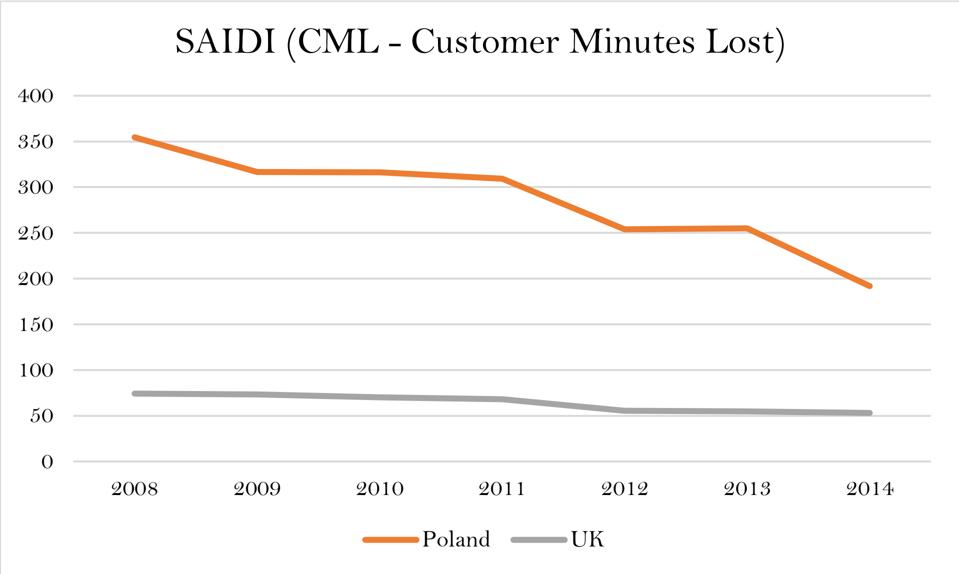
Glossary of terms

Free disposability:	We can produce less with more: (x, y) ∈ T, x' ≥ x and y' ≥ y => (x', y') ∈ T. Convexity: Any weighted average of feasible production plans is feasible as well: (x, y) ∈ T, (x', y') ∈ T, α ∈ [0,1] => α (x, y) + (1 - α) (x', y') ∈ T.
γ - returns to scale	Production can be scaled with any of a given set of factors: (x, y) ∈ T; κ ∈ Γ(γ) => κ* (x, y) ∈ T.

Figures

Figure 1

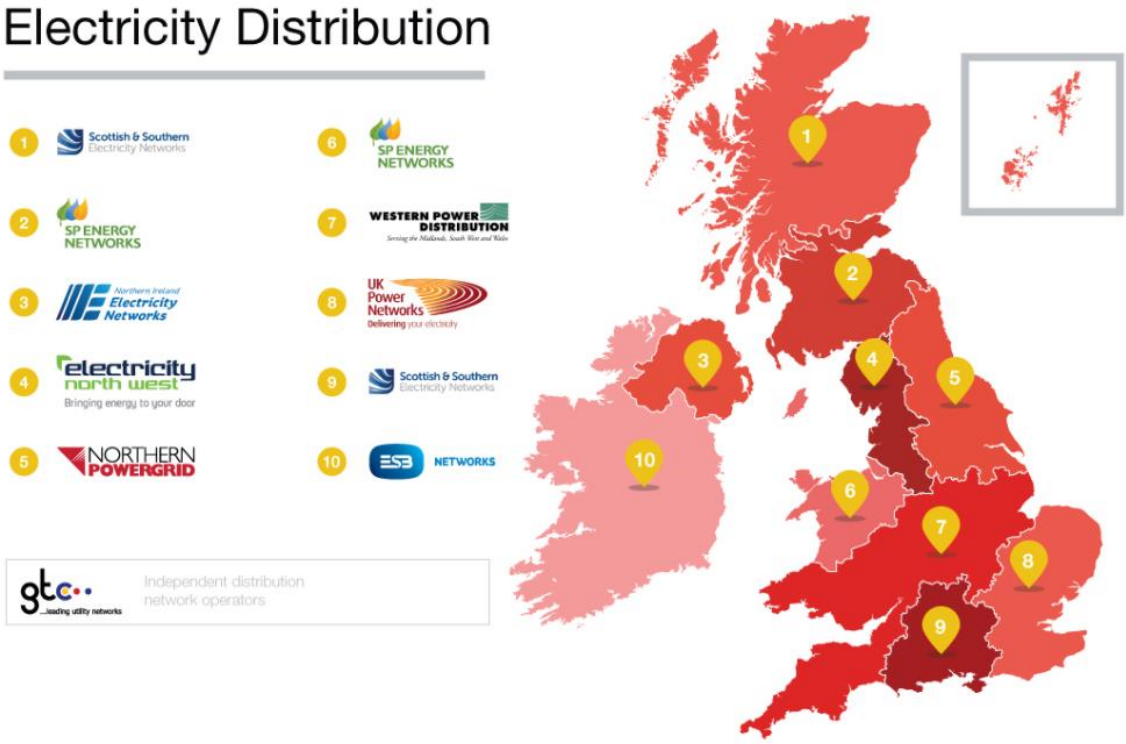
SAIDI comparison UK and Poland (2008-2014)



Source: Author

Figure 2

UK DSO geographical coverage



Source: Energy Network Association (n.d.)

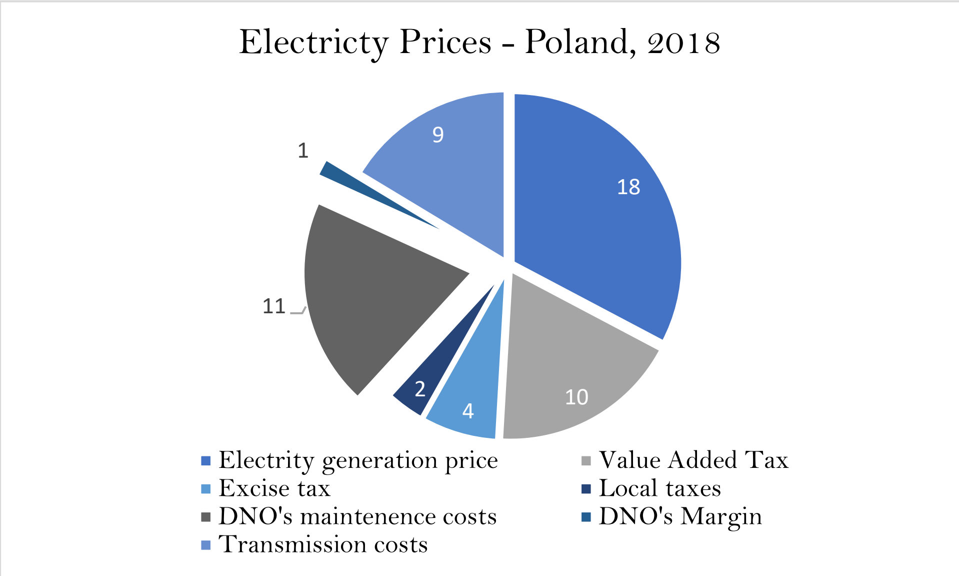
Figure 3

Territorial coverage of the five main Polish DSOs

Source: PTPiREE (2017)

Figure 4

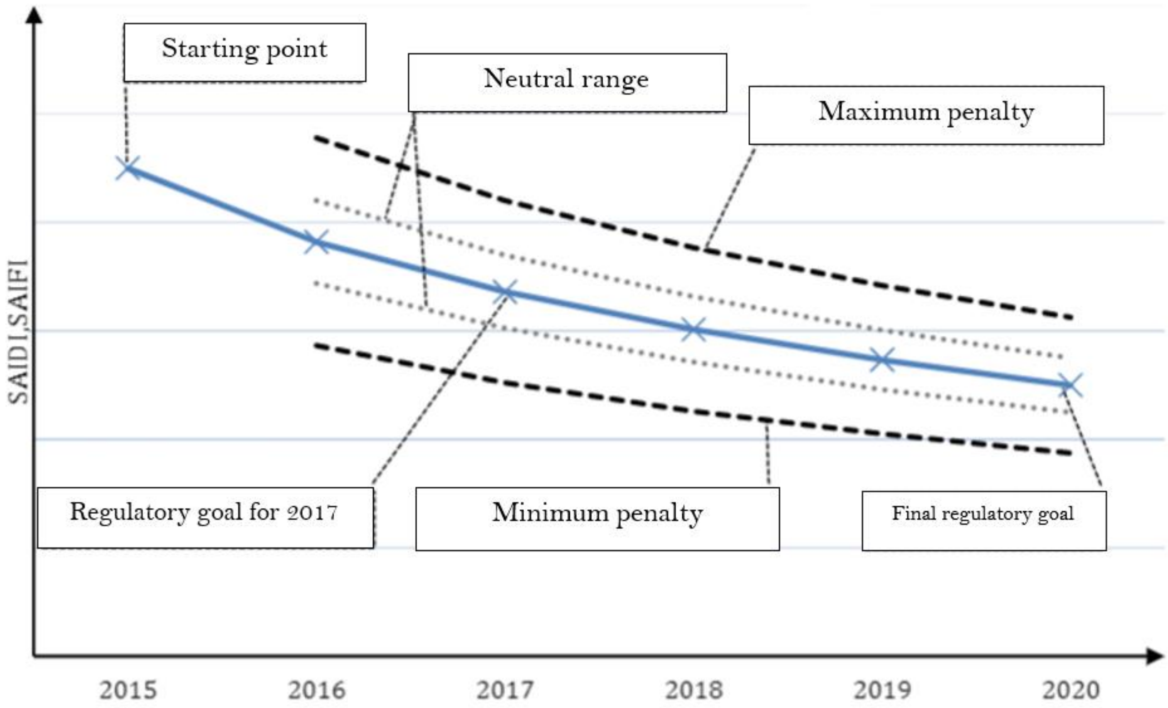
Final Electricity Prices in Poland, 2018 (gr/kW)



Source: Cena Pradu (n.d.)

Figure 5

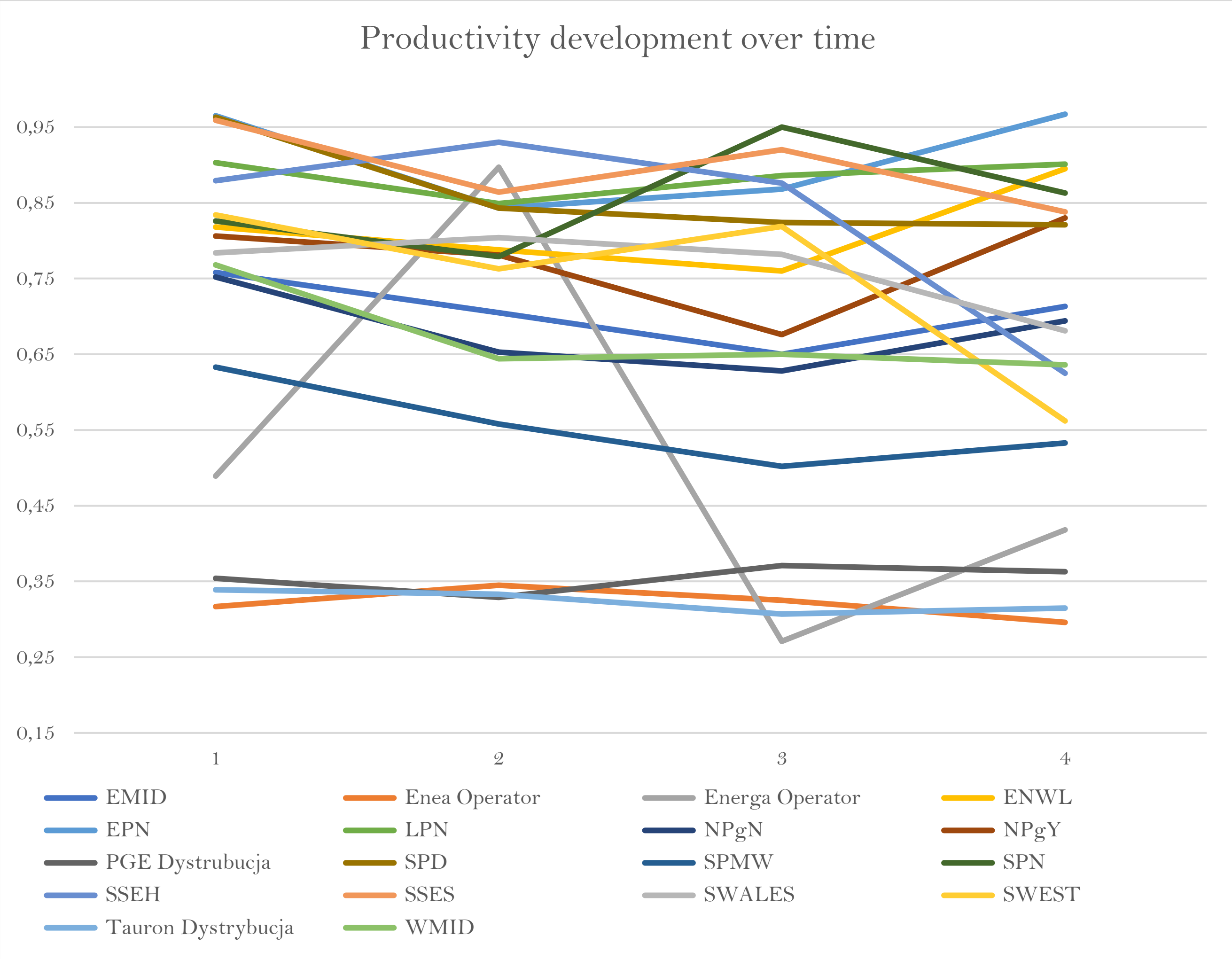
URE scheme for SAIDI/SAIFI regulated levels



Adapted from: URE, 2015b

Figure 6

Productivity development over 2013-2015 (M3)



Source: Author

Tables

Table 1

Individual performances ($\widehat{\theta^k}$) in the eight DEA models

	Model1	Model2	Model3	Model4	Model5	Model6	Model7	Model8
EMID1	0.620	0.708	0.758	0.753	0.753	0.697	0.753	0.752
EMID2	0.576	0.653	0.705	0.700	0.719	0.642	0.701	0.699
EMID3	0.531	0.601	0.650	0.644	0.675	0.586	0.646	0.644
EMID4	0.577	0.672	0.713	0.706	0.754	0.662	0.706	0.705
EneaO1	0.121	0.197	0.317	0.316	0.316	0.196	0.324	0.323
EneaO2	0.131	0.211	0.345	0.344	0.344	0.210	0.354	0.353
EneaO3	0.122	0.198	0.325	0.325	0.325	0.197	0.333	0.332
EneaO4	0.125	0.194	0.296	0.295	0.295	0.193	0.301	0.300

EnergaO1	0.156	0.258	0.489	0.487	0.487	0.258	0.621	0.616
EnergaO2	0.305	0.498	0.897	0.893	0.893	0.497	0.895	0.887
EnergaO3	0.095	0.152	0.271	0.270	0.270	0.151	0.280	0.277
EnergaO4	0.147	0.234	0.418	0.416	0.416	0.233	0.449	0.445
ENWL1	0.683	0.806	0.818	0.814	0.813	0.801	0.823	0.821
ENWL2	0.651	0.779	0.788	0.783	0.781	0.773	0.792	0.791
ENWL3	0.638	0.744	0.760	0.752	0.747	0.733	0.764	0.762
ENWL4	0.724	0.890	0.895	0.928	0.924	0.871	0.898	0.897
EPN1	0.783	0.927	0.965	0.962	0.957	0.923	0.857	0.846
EPN2	0.670	0.804	0.843	0.841	0.837	0.801	0.791	0.783
EPN3	0.689	0.828	0.868	0.865	0.862	0.825	0.851	0.842
EPN4	0.726	0.915	0.967	0.964	0.960	0.909	0.888	0.879
LPN1	0.975	0.901	0.903	0.892	0.896	0.890	0.917	0.911
LPN2	0.898	0.848	0.849	0.838	0.840	0.840	0.866	0.861
LPN3	0.944	0.884	0.886	0.880	0.881	0.894	0.902	0.897
LPN4	0.915	0.906	0.901	0.828	0.825	0.858	0.921	0.916
NPgN1	0.613	0.731	0.752	0.747	0.745	0.726	0.847	0.757
NPgN2	0.525	0.635	0.653	0.656	0.654	0.630	0.743	0.657
NPgN3	0.505	0.610	0.628	0.626	0.621	0.602	0.712	0.631
NPgN4	0.543	0.668	0.694	0.712	0.709	0.662	0.788	0.698
NPgY1	0.683	0.801	0.806	0.805	0.805	0.797	0.809	0.810
NPgY2	0.653	0.781	0.781	0.779	0.779	0.777	0.784	0.785
NPgY3	0.564	0.676	0.676	0.674	0.673	0.671	0.678	0.679
NPgY4	0.690	0.832	0.830	0.824	0.816	0.820	0.834	0.835
PGE	0.114	0.215	0.354	0.353	0.353	0.215	0.875	0.871
PGE	0.108	0.200	0.329	0.328	0.328	0.199	0.810	0.806
PGE	0.124	0.226	0.371	0.371	0.371	0.225	0.892	0.884
PGE	0.124	0.221	0.363	0.363	0.363	0.220	0.821	0.814
SPD1	0.771	0.901	0.963	0.948	0.940	0.892	0.955	0.964
SPD2	0.676	0.791	0.843	0.832	0.831	0.781	0.838	0.845
SPD3	0.660	0.773	0.824	0.827	0.856	0.755	0.818	0.825
SPD4	0.595	0.743	0.821	0.834	0.860	0.754	0.814	0.820
SPMW1	0.494	0.537	0.633	0.663	0.653	0.529	0.651	0.634
SPMW2	0.445	0.485	0.558	0.569	0.562	0.478	0.585	0.560
SPMW3	0.400	0.438	0.502	0.529	0.519	0.425	0.527	0.504
SPMW4	0.415	0.475	0.533	0.572	0.565	0.467	0.567	0.535
SPN1	0.682	0.840	0.826	0.822	0.821	0.833	0.830	0.831
SPN2	0.639	0.793	0.779	0.776	0.776	0.787	0.783	0.784
SPN3	0.784	0.967	0.950	0.942	0.922	0.946	0.953	0.956
SPN4	0.653	0.881	0.863	0.860	0.845	0.860	0.868	0.869
SSEH1	0.488	0.504	0.879	0.884	0.888	0.485	0.882	0.879
SSEH2	0.514	0.534	0.930	0.904	0.900	0.524	0.926	0.929
SSEH3	0.480	0.501	0.876	0.852	0.859	0.490	0.875	0.876
SSEH4	0.312	0.355	0.625	0.624	0.645	0.344	0.622	0.623
SSES1	0.834	0.869	0.959	0.957	0.957	0.863	0.935	0.928
SSES2	0.753	0.789	0.864	0.862	0.862	0.784	0.845	0.839
SSES3	0.800	0.842	0.920	0.918	0.917	0.837	0.898	0.892
SSES4	0.701	0.812	0.838	0.837	0.835	0.807	0.815	0.808
SWALES1	0.612	0.678	0.784	0.827	0.895	0.877	0.933	0.788
SWALES2	0.627	0.697	0.804	0.847	0.897	0.864	0.954	0.808
SWALES3	0.609	0.678	0.782	0.806	0.890	0.856	0.928	0.786

SWALES4	0.520	0.604	0.681	0.744	0.894	0.880	0.833	0.684
SWEST1	0.554	0.705	0.834	0.817	0.829	0.699	0.838	0.832
SWEST2	0.507	0.646	0.763	0.767	0.779	0.650	0.767	0.761
SWEST3	0.543	0.693	0.819	0.825	0.842	0.700	0.822	0.817
SWEST4	0.359	0.483	0.562	0.566	0.577	0.486	0.569	0.561
TauronD1	0.174	0.221	0.339	0.338	0.337	0.221	0.900	0.894
TauronD2	0.171	0.220	0.333	0.332	0.332	0.219	0.882	0.874
TauronD3	0.158	0.200	0.307	0.305	0.305	0.199	0.823	0.810
TauronD4	0.164	0.203	0.315	0.314	0.313	0.203	0.823	0.811
WMID1	0.628	0.742	0.768	0.766	0.765	0.737	0.770	0.766
WMID2	0.528	0.620	0.644	0.643	0.641	0.614	0.647	0.643
WMID3	0.532	0.625	0.650	0.648	0.644	0.616	0.652	0.649
WMID4	0.516	0.617	0.636	0.634	0.636	0.605	0.636	0.634
Note: Green for lowest items, Red for highest items, Yellow for the highest and lowest overall performers.								

Source: Author

Table 2

Translog and Cobb-Douglas models: Likelihood Ratio Test

Likelihood ratio test (LR): Model Translog and Model Cobb-Douglas	
LR chi2(8) = -33.34	Prob > chi2 = 1.0000

Source: Author

Table 3

SFA Model 2 results3

ln(TOTEX)		Coef.	Std. Err.	z	P>|z|	95% C.I.	
Frontier β	ln(CUST)	0.335	0.317	1.06	0.290	-0.286	0.956
	ln(LENGTH)	0.315	0.144	2.19	0.029	0.032	0.598
	Ln(UNITS)	0.187	0.289	0.65	0.518	-0.380	0.752
	(Intercept)	-3.343	3.278	-1.02	0.308	-9.765	-3.079
	D_{PL}(COUNTRY)	1.067	0.169	6.30	0.000*	0.735	1.399
Note: Values marked with an asterisk* are significant at 5% level.							

Source: Author

Table 4

Individual efficiency scores in the SFA Model 3 4

DMU	Year	Lower bound (95%)	Efficiency score	Upper bound (95%)
EMID	2013	0.703677	0.802827	0.915609
EMID	2014	0.703677	0.802827	0.915609
EMID	2015	0.703677	0.802827	0.915609
EMID	2016	0.703677	0.802827	0.915609

EneaO	2013	0.734531	0.837505	0.952065
EneaO	2014	0.734531	0.837505	0.952065
EneaO	2015	0.734531	0.837505	0.952065
EneaO	2016	0.734531	0.837505	0.952065
EnergaO	2013	0.894265	0.965128	0.998882
EnergaO	2014	0.894265	0.965128	0.998882
EnergaO	2015	0.894265	0.965128	0.998882
EnergaO	2016	0.894265	0.965128	0.998882
ENWL	2013	0.822185	0.922363	0.995256
ENWL	2014	0.822185	0.922363	0.995256
ENWL	2015	0.822185	0.922363	0.995256
ENWL	2016	0.822185	0.922363	0.995256
EPN	2013	0.844356	0.938109	0.997111
EPN	2014	0.844356	0.938109	0.997111
EPN	2015	0.844356	0.938109	0.997111
EPN	2016	0.844356	0.938109	0.997111
LPN	2013	0.80852	0.911372	0.99335
LPN	2014	0.80852	0.911372	0.99335
LPN	2015	0.80852	0.911372	0.99335
LPN	2016	0.80852	0.911372	0.99335
NPgN	2013	0.758469	0.863456	0.973555
NPgN	2014	0.758469	0.863456	0.973555
NPgN	2015	0.758469	0.863456	0.973555
NPgN	2016	0.758469	0.863456	0.973555
NPgY	2013	0.793778	0.898429	0.990164
NPgY	2014	0.793778	0.898429	0.990164
NPgY	2015	0.793778	0.898429	0.990164
NPgY	2016	0.793778	0.898429	0.990164
PGE	2013	0.703952	0.80314	0.915956
PGE	2014	0.703952	0.80314	0.915956
PGE	2015	0.703952	0.80314	0.915956
PGE	2016	0.703952	0.80314	0.915956
SPD	2013	0.863762	0.949881	0.998043
SPD	2014	0.863762	0.949881	0.998043
SPD	2015	0.863762	0.949881	0.998043
SPD	2016	0.863762	0.949881	0.998043
SPMW	2013	0.605974	0.691438	0.788956

SPMW	2014	0.605974	0.691438	0.788956
SPMW	2015	0.605974	0.691438	0.788956
SPMW	2016	0.605974	0.691438	0.788956
SPN	2013	0.870208	0.953412	0.998268
SPN	2014	0.870208	0.953412	0.998268
SPN	2015	0.870208	0.953412	0.998268
SPN	2016	0.870208	0.953412	0.998268
SSEH	2013	0.780118	0.8855	0.985594
SSEH	2014	0.780118	0.8855	0.985594
SSEH	2015	0.780118	0.8855	0.985594
SSEH	2016	0.780118	0.8855	0.985594
SSES	2013	0.815341	0.916981	0.9944
SSES	2014	0.815341	0.916981	0.9944
SSES	2015	0.815341	0.916981	0.9944
SSES	2016	0.815341	0.916981	0.9944
SWALES	2013	0.828284	0.92695	0.995886
SWALES	2014	0.828284	0.92695	0.995886
SWALES	2015	0.828284	0.92695	0.995886
SWALES	2016	0.828284	0.92695	0.995886
SWEST	2013	0.771507	0.876937	0.981608
SWEST	2014	0.771507	0.876937	0.981608
SWEST	2015	0.771507	0.876937	0.981608
SWEST	2016	0.771507	0.876937	0.981608
TauronD	2013	0.676687	0.772113	0.880963
TauronD	2014	0.676687	0.772113	0.880963
TauronD	2015	0.676687	0.772113	0.880963
TauronD	2016	0.676687	0.772113	0.880963
WMID	2013	0.696592	0.79478	0.906618
WMID	2014	0.696592	0.79478	0.906618
WMID	2015	0.696592	0.79478	0.906618
WMID	2016	0.696592	0.79478	0.906618

Source: Author

Table 5

Results from the Bayesian SFA (Half- Normal; Persistent Effects)

CAME	1,197879	
	Coefficient	Error
CUST	0.556634	0.08997
Length	0.300142	0.1182
CountryPL	0.762415	0.160252
QualityPL	0.35918	0.136228
Const	-5.90482	0.955141

Source: Adapted from Makieła (2018)

Table 6

Individual efficiencies form the Bayesian SFA (Half- Normal; Persistent Effects) and Varying Effects Model

DMU	**Year**	**VED**		**Bayesian SFA**	
		Efficiency scores	**Error**	**Efficiency scores**	**Error**
EMID	2013	0.8450	0.0699	0.806625	0.086006
ENWL	2013	0.9197	0.0503	0.86584	0.077033
EPN	2013	0.9623	0.0330	0.909311	0.065492
LPN	2013	0.9142	0.0604	0.864559	0.081097
NPgN	2013	0.9171	0.0572	0.856977	0.079506
NPgY	2013	0.9199	0.0529	0.860096	0.07843
SPD	2013	0.9745	0.0253	0.928909	0.056273
SPMW	2013	0.7515	0.0709	0.719282	0.087294
SPN	2013	0.9425	0.0429	0.89068	0.071184
SSEH	2013	0.9416	0.0450	0.850328	0.087459
SSES	2013	0.9525	0.0391	0.88812	0.071931
SWALES	2013	0.9300	0.0498	0.869171	0.077798
SWEST	2013	0.9374	0.0448	0.879683	0.073939
WMID	2013	0.8854	0.0581	0.8319	0.082699
Enea Operator	2013	0.3226	0.0730	0.639502	0.115081
Energa Operator	2013	0.4357	0.0873	0.807743	0.109782
PGE Dystrubucja	2013	0.3139	0.0782	0.633919	0.115493
Tauron Dystrybucja	2013	0.3197	0.0639	0.634663	0.114938
EMID	2014	0.7879	0.0783	0.760403	0.087775
ENWL	2014	0.9370	0.0503	0.852298	0.079668

EPN	2014	0.9091	0.0549	0.85425	0.081576
LPN	2014	0.8703	0.0701	0.83248	0.086652
NPgN	2014	0.8280	0.0732	0.784364	0.087501
NPgY	2014	0.9403	0.0493	0.850716	0.080392
SPD	2014	0.9480	0.0393	0.889317	0.070617
SPMW	2014	0.6736	0.0700	0.654377	0.083576
SPN	2014	0.9171	0.0533	0.866202	0.076884
SSEH	2014	0.9598	0.0355	0.87535	0.080687
SSES	2014	0.9099	0.0558	0.842258	0.082147
SWALES	2014	0.9415	0.0448	0.880405	0.075027
SWEST	2014	0.8858	0.0602	0.837831	0.082129
WMID	2014	0.7554	0.0710	0.729019	0.087005
Enea Operator	2014	0.3408	0.0669	0.678891	0.11709
Energa Operator	2014	0.7967	0.1089	0.96899	0.029514
PGE Dystrubucja	2014	0.2902	0.0710	0.595945	0.111855
Tauron Dystrybucja	2014	0.3113	0.0581	0.627783	0.11425
EMID	2015	0.7259	0.0819	0.712723	0.087083
ENWL	2015	0.9087	0.0620	0.82395	0.083667
EPN	2015	0.9197	0.0523	0.868224	0.078393
LPN	2015	0.8927	0.0672	0.853578	0.083603
NPgN	2015	0.7964	0.0721	0.760417	0.087852
NPgY	2015	0.8575	0.0720	0.770561	0.086857
SPD	2015	0.9352	0.0455	0.879538	0.073313
SPMW	2015	0.6027	0.0677	0.599822	0.078834
SPN	2015	0.9723	0.0271	0.931392	0.055407
SSEH	2015	0.9400	0.0456	0.847355	0.088016
SSES	2015	0.9383	0.0455	0.873846	0.075764
SWALES	2015	0.9309	0.0493	0.869216	0.077834
SWEST	2015	0.9252	0.0493	0.871859	0.075862
WMID	2015	0.7589	0.0720	0.734181	0.086955
Enea Operator	2015	0.3250	0.0687	0.846058	0.087706
Energa Operator	2015	0.2491	0.0498	0.704414	0.094451
PGE Dystrubucja	2015	0.3277	0.0768	0.85683	0.08531
Tauron Dystrybucja	2015	0.2871	0.0564	0.778437	0.095828
EMID	2016	0.8037	0.0812	0.780349	0.08773
ENWL	2016	0.9572	0.0356	0.91166	0.063913
EPN	2016	0.9587	0.0349	0.910669	0.065149

LPN	2016	0.8933	0.0671	0.86127	0.081773
NPgN	2016	0.8609	0.0652	0.819508	0.084755
NPgY	2016	0.9333	0.0458	0.88283	0.073012
SPD	2016	0.9198	0.0520	0.874284	0.074947
SPMW	2016	0.6549	0.0696	0.645232	0.082632
SPN	2016	0.9493	0.0402	0.905145	0.066484
SSEH	2016	0.7056	0.0798	0.656655	0.094385
SSES	2016	0.9056	0.0547	0.85336	0.080196
SWALES	2016	0.8529	0.0691	0.811303	0.087225
SWEST	2016	0.6696	0.0748	0.665022	0.084198
WMID	2016	0.7470	0.0729	0.729376	0.08728
Enea Operator	2016	0.2985	0.0533	0.81709	0.093413
Energa Operator	2016	0.3791	0.0692	0.915194	0.064999
PGE Dystrubucja	2016	0.3182	0.0710	0.84599	0.087546
Tauron Dystrybucja	2016	0.2868	0.0488	0.792102	0.094817

Source: Adapted from Makieła (2018)

Equations

Equation 1

Return on capital employed: Poland

$$Zt \ = (r + \delta) = \ WRAt \ * \ WACCt \ * \ Qt \ * \ WRt$$

Zt – return on capital employed for the year t.

RVAt – regulatory value of assets for year t

WACCt – weighted average cost of capital set for year t

Qt – general coefficient of implementation of the quality regulation ($1< Qt < 0.85$)

WRt – regulatory coefficient, set individually for DSOs, reflecting their innovation efforts

Source: URE, 2015a