Deloitte.

Annex 4 - GD17 Efficiency Advice

Relative efficiency of Northern Ireland Gas Distribution Networks

Final report

11 March 2016

Important Notice from Deloitte

This final report (the "Final Report") has been prepared by Deloitte LLP ("Deloitte") for the Northern Ireland Authority for Utility Regulation in accordance with the contract with them dated November 30th, 2015 ("the Contract") and on the basis of the scope and limitations set out below.

The Final Report has been prepared solely for the purposes of the estimation of relative efficiency of Northern Ireland gas distribution networks, as set out in the Contract. It should not be used for any other purpose or in any other context, and Deloitte accepts no responsibility for its use in either regard.

The Final Report is provided exclusively for the Northern Ireland Authority for Utility Regulation's use under the terms of the Contract. No party other than the Northern Ireland Authority for Utility Regulation's is entitled to rely on the Final Report for any purpose whatsoever and Deloitte accepts no responsibility or liability or duty of care to any party other than the Northern Ireland Authority for Utility Regulation in respect of the Final Report or any of its contents. .

The information contained in the Final Report has been obtained from the Northern Ireland Authority for Utility Regulation and third party sources that are clearly referenced in the appropriate sections of the Final Report. Deloitte has neither sought to corroborate this information nor to review its overall reasonableness. Further, any results from the analysis contained in the Final Report are reliant on the information available at the time of writing the Final Report and should not be relied upon in subsequent periods.

Contents

Executive summary

In order to inform the GD17 price control process for the Northern Ireland (NI) gas distribution networks (GDNs), the Northern Ireland Utility Regulator has commissioned Deloitte to undertake an econometric analysis to estimate the efficiency of the operational expenditure of Firmus Energy Distribution Ltd (FE) and Phoenix Natural Gas Ltd (PNGL). Due to the lack of sufficient number of comparators, data for the eight Great Britain (GB) GDNs have been used to benchmark the NI GDNs. There are a number of particular challenges in applying econometric benchmarking analysis across the GB and NI GDN sample.

- The GB GDNs all operate on a scale materially greater than either FE or PNGL. Whilst the econometric analysis seeks to allow for economies of scale the extent to which this is fully captured is challenging as the dataset is dominated by larger GDNs. As a result the impact of scale on costs, and subsequently relative efficiency, may be over- or under-estimated.
- FE has a significantly different profile to any of the other GDNs in terms of the ratios between network length, customer numbers and volume of gas supplies. This results in significant challenges in assessing the extent to which FE costs are inefficient or are due to the characteristics of the business. As such, FEs relative efficiency has been computed by estimating a model using the GB only or GB and PNGL data and fitting the model to the FE data. A detailed analysis of special factors driving cost differences between FE and other GDNs, which is outside of the scope of this report, would be required to isolate these effects.
- Finally, the age and pipeline materials for the NI GDNs are different to the GB GDNs which may influence relative costs.

The key findings are that:

- The main driver of cost variation over time and across GDNs is scale. The models estimate that a 1% increase in the scale of a GDN is expected to increase costs by 0.69%-0.81%.
- There is some indication that scale effects are more pronounced in NI, potentially due to smaller scale and greater economies of scale realised when starting from a lower base.
- Other variables such as network composition and network length per customer are statistically insignificant. This may reflect the non-existence of a relationship or the small sample size used in the estimation and/or limited data variability, which might hinder models' ability to identify significant effects.
- The relative efficiency estimates, in particular for FE, are quite sensitive to the year used as the basis for the computation of efficiency. The results are also sensitive to the sample used for the estimation and whether the age or state of networks is accounted for, or not. The 2014 FE efficiency ranges from 5.1% to 8.7% when both GB and PNGL data are used in the estimation and between 0% and 29% when GB only data are considered. For PNGL, the range of efficiencies is 5.8%-8.1% and 2.6%-25% in the GB and PNGL, and GB only samples, respectively.

In order to mitigate estimation uncertainty and the challenges associated with the comparability of the NI and GB GDNs, the efficiency determination should consider combining the econometric analysis with special factor adjustments (both positive and negative) and/or additional information such as the bottom-up model UR is developing.

1 Introduction

Within Northern Ireland (NI) there are currently three monopoly Gas Distribution Network (GDN) companies: Phoenix Natural Gas Ltd (PNGL), Firmus Energy Distribution Ltd (FE) and Scotia Gas Networks Northern Ireland Ltd (SGN). These companies are each the sole provider of gas distribution services through their network within each of their respective geographical license areas. As local utility monopolies they are regulated by the Northern Ireland Authority for Utility Regulation (UR) through regular price controls.

The next price control, GD17, will run for a period of 6 years from 2017 to 2022. The key objective of the price control is to determine GDNs' relative efficiency and cost allowances and subsequently the maximum prices they can charge to consumers. Ultimately, like other price controls, the aim is to ensure that resources are allocated efficiently and prices consumers pay are as low as they ought be.

1.1 The GD17 process

On 19 December 2014, the Utility Regulator published a discussion document[1] setting out their proposed basis for determining the allowed distribution charges under the GD17 price control and invited feedback on the proposed approach.

Following feedback, on 17 April 2015 the UR published an update on their approach[2] which sets out the key considerations and aspects of the GD17 approach including:

- **Catch-up efficiency.** The UR stated that top-down econometrics and bottom-up approaches will be considered for the estimation of the relative efficiency of NI GDNs;
- **Frontier shift.** Expected productivity changes and real price effects over the price control period will also be taken into consideration for the determination of the expenditure allowances;
- **Comparators.** Due to the small number of operators in NI, the UR indicated that it will use data from Great Britain (GB) GDNs to benchmark the NI networks; and
- **Comparability and special factors**. To ensure a like-for-like comparison of NI GDNs to their GB counterpart's, the UR stated that it will consider which non-controllable or controllable costs might be excluded to improve comparability.

1.2 This report

This report describes the econometric analysis undertaken by the Utility Regulator and Deloitte to inform the GD17 process. In particular, a number of econometric techniques have been considered and applied to a dataset comprised of NI and GB GDN data with the objective of obtaining estimates of the catch-up efficiency for NI GDNs.

This remainder of this report is structured as follows:

- Section 2 describes the econometric methodology and key challenges;
- Section 3 describes the data; and
- Section 4 sets out the econometric results.

[1] http://www.uregni.gov.uk/uploads/publications/2014-12-19_GD17_Price_Control_Scope_-_Final.pdf
[2] http://www.uregni.gov.uk/uploads/publications/2015-04-17_GD17_-_Approach_Document_-_Final.pdf

2 Methodology

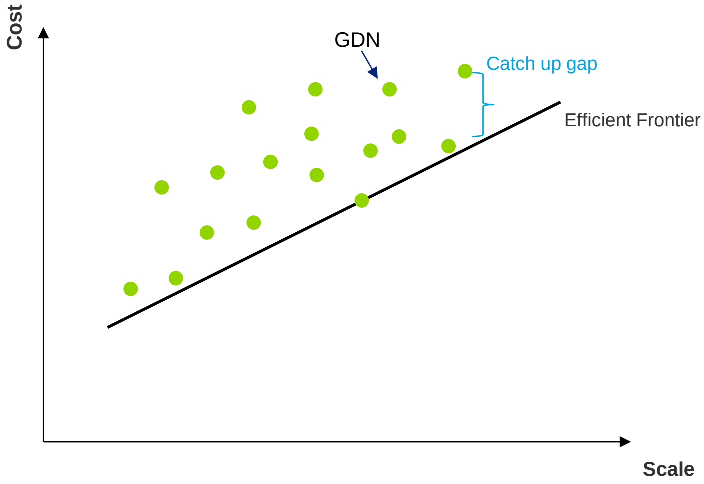
Efficiency benchmarking is carried out by comparing the costs to different companies of delivering a product or service to customers; in the case of GDNs, the cost of developing, operating and maintaining the medium and low pressure gas network. GDNs operate in different environments and are subject to a number of exogenous factors that may lead to differences in operating costs, for example number of customers served, total volume supplied, network length, regional labour costs and network age. If this heterogeneity between GDNs is not taken into account, differences in operating costs cannot be attributed to differences in efficiency. The primary challenge in measuring efficiency is therefore to control for these exogenous factors.

Econometrics has been widely applied in previous price controls, both in GB and NI.[3] Its appeal primarily stems from the ability to control for multiple exogenous factors in relative efficiency estimation.

2.1 Catch-up efficiency

The objective of the analysis is to estimate the efficient frontier as illustrated in Figure 1. The frontier demonstrates the minimum cost or resource that is required to distribute different gas volumes or service different numbers of customers.[4] Catch-up efficiency measures the efficiency savings that a GDN could achieve if it were to become as efficient as the most efficient GDN.

Figure 1: Efficient frontier



The catch-up efficiency is used by a regulator to inform the decision regarding the allowed increase (or decrease) in cost for a regulated utility, based on the principle that costs should be fully recovered by an efficient company. Typically, a particular point in the efficiency distribution is used as the point at which a company is considered efficient. Upper-quartile and upper-decile are commonly used. Recognising that there is uncertainty around any efficiency measure, and that improving efficiency takes time, a common approach is to use a combination of a non-frontier efficiency benchmark and a transition period to achieve the benchmark efficiency in determining the profile of allowed revenue.

[3] For instance, see "RIIO-GD1: Final Proposals - Supporting document - Cost efficiency. Ofgem"; "Draft determinations for PR14. Ofwat (2014)"; "NIE T&D RP5 Opex Evaluation, Efficiency Assessment of NIE's operating expenditure. UR".
[4] The frontier depicted in this example is linear and assumes that only scale drives cost differentials between GDNs.

2.2 Comparators

Econometric models use data variability to separate that component of costs which could be attributable to management decisions (efficiency) and costs which reflect exogenously determined operating conditions. The ability to separate these components depends inter alia on the sample size and the variation in costs and cost drivers across comparators and over time.

There are three GDNs operating in NI and only two for which historical data are available (PNGL and FE). SGN was granted a licence in early 2015 and therefore no sufficient data for this GDN are available to include in the analysis.

In order to increase the sample size, the UR has obtained data from GB GDNs and proposed to benchmark PNGL and FE against their GB counterparts.[5] Benchmarking companies across different countries is not uncommon in regulation. For instance, due to the lack of comparators in England, in the 2009 Business Connectivity Market Review BT was benchmarked against the US Local Exchange Carriers.[6] The Office of Rail Regulation has also collected and used data on international comparators.[7]

Although using comparators from different regions[8] is an effective way to deal with small sample sizes, it introduces another challenge, in particular around comparability. This issue is discussed in detail in section 2.6.

2.3 Econometric methodologies

There are three commonly used econometric methodologies in efficiency analysis.

- **Pooled OLS (POLS).** The main advantage of POLS is that it is simple to apply. The main drawback is that it does not separate data noise from efficiency, and for this reason, has the tendency to over-estimate the efficiency gap. Given that the estimator does not take into account the panel structure of the dataset and the fact that observations for a specific operator are correlated over time, standard errors are generally biased downwards. However, this can be addressed by using cluster standard errors.

- **Panel random effects (RE).** The RE estimator is designed for panel datasets and under specific assumptions provides more efficient coefficient estimates compared with POLS.[9] For instance, it requires that each of the two components of the composite error are identical and independently distributed. If these assumptions are violated, one might utilise clustered standard errors along with RE.[10]

- **Stochastic frontier analysis (SFA).** SFA is specifically designed to estimate comparative efficiency. SFA models are estimated by maximum likelihood and apply distributional assumptions to decompose the model error into a random component and an efficiency

---

[5] Another way to increase the sample size is to use panel data, i.e. cross-sectional (GDNs) together with historic time-series data, which is the approach used in this study.

[6] The sample included 26 US LECs over the period from 1996 to 2006. http://stakeholders.ofcom.org.uk/binaries/consultations/llcc/responses/BT2.pdf

[7] The Office of Rail Regulation used 12 European rail infrastructure managers operating in the following countries: UK, Austria, Belgium, Denmark, Finland, Germany, Ireland, Italy, the Netherlands, Norway, Sweden and Switzerland. [7] http://orr.gov.uk/__data/assets/pdf_file/0015/4713/econometric_update_2010_orr_benchmarking_report.pdf

[8] Whilst NI and GB regions are subject to slightly different regulatory arrangements, the two regions operate within the same national economy and use pound sterling as the currency.

[9] RE is best described relative to POLS. In a panel setting, data variation has two components: within variation denotes variation in costs around the average costs for each GDN; between variation considers the variation around the set of mean GDN costs. Whereas POLS gives equal weight to these two components of variation, the RE estimator weights these components inversely proportional to their variance. In this sense, the RE estimator can also be interpreted as a Generalised Least Squares estimator.

[10] In theory, Fixed Effects (FE) could be an alternative econometric approach. However, it is typically not used in the context of efficiency benchmarking due to the nature of the underlying data. When explanatory variables are slow moving, they tend to be highly correlated with the fixed effects. This makes it difficult to isolate efficiency captured by the fixed effects from scale and other cost drivers.

component. The simplest version of SFA assumes that efficiency remains constant over time. More sophisticated models allow for time-varying efficiency, the efficiency factor to be modelled as a function of exogenous variables, and fixed effects that distinguish between company specific efficiency and unobserved heterogeneity (i.e. company specific effects that are unrelated to efficiency). SFA models are more complex and require larger datasets than POLS and RE and are typically applied in situations when the number of comparators is significantly larger than the current sample.

Figure 2: Econometric model

In its simplest form, econometric analysis is used to investigate how the 'dependent' variable of interest y (e.g. cost) varies as a function of a set of *n* independent variables: X_1, X_2,..., X_n. This relationship can be specified as:

$$y_{i,t} = \alpha + \beta_1 X_{1i,t} + \beta_2 X_{2i,t} + \ldots\ \beta_n X_{ni,t} + \varepsilon_{i,t}$$

Where:

- Each β 'coefficient' represents the estimated relationship that the associated explanatory variable holds with the dependent variable y. For example, β_1 shows the estimated change in y resulting from a one unit change in X_1, all else being equal.[11]
- ε denotes a random error term, namely the vertical distance that a particular observation lies from the linear relationship estimated.
- α represents the intercept.

This equation assumes that observations across GDNs (i) and over time (t) are available to estimate the parameters of interest.

The model predictions provide an estimate of efficient costs for each GDN given these determinants. The difference $(\varepsilon_{i,t})$ between the cost predicted by the model and the observed cost is then used to infer the relative efficiency of each operating unit.

2.4 Core methodology for GD17

For the GD17 price control, the core elements of the analysis have been determined as:

1. **Cost measure.** The focus of the efficiency analysis will be operating expenditure (OPEX).
2. **Model specification.** Based on a review of precedent in the price regulation of GDNs, three key variables are considered: (1) scale; (2) network composition; and (3) regional wages. Time-specific effects that are common across all GDNs are also controlled for through time dummy variables.[12] The choice of control variables has been determined on the basis of the type of factors that are expected to drive costs and which are outside management's control, data availability, and the number of comparators.
3. **Estimation technique.** Given the small number of comparators and following regulatory precedence[13], this study primarily uses POLS. However, the sensitivity of the results are examined by also considering RE and SFA models.

[11] If each variable is specified in natural logarithm terms, this can be interpreted as the percentage change in y_i given a 1% change in X_{1i}.
[12] Time effects are often included in regression models to control for unobserved factors that are common across cross-section units (i.e. GDNs) but vary over time.
[13] For instance, see Ofgem RIIO-GD1 and Ofgem RIIO-ED1.

2.5 Composite Scale Variable

The key factors that drive cost differentials between GDNs are the network size, number of customers and gas volume. The main challenge in including these factors within an econometric analysis is that the variables are highly correlated and hence only one or a subset of them could be included in the model.

Ofgem in RIIO-ED1 and RIIO-GD1, as well as the UR in RP5, PC13 and PC15, used a composite scale variable (CSV). CSVs are used to address the problem of multicollinearity in a multiple regression model. Multicollinearity occurs when explanatory variables are highly correlated and the sample size is not large enough to isolate their individual impact. A CSV effectively combines two or more variables into one explanatory factor of scale, using a weighted index, which allows capturing the impact of all variables and avoiding the problem of multicollinearity. Following Ofgem and precedence in the NI regulation, scale is primarily captured through a CSV.

2.6 Key challenges and mitigations

The overall framework described in the previous section is consistent with a variety of previous regulatory processes and practices. However, any efficiency analysis of the Northern Ireland GDNs has some particular challenges primarily associated with benchmarking companies from different regions, the small number of comparators and their varying stage of development.

Scope of activities

The Regulatory Instructions and Guidance for GDNs in NI and GB are very similar. Therefore the categories under which GDNs report costs and how they allocate costs between these categories are similar in many ways. However, there are differences in the scope of activities undertaken by GDNs in GB and NI and differences in legislation that applies in GB and NI.

In order to conduct a more like-for-like comparison between GB and NI GDNs, the UR, in consultation with Deloitte, has made a number of exclusions from the cost data for both NI and GB GDNs to remove cost items from the analysis where GDNs in one region do not face equivalent responsibilities or costs to those in the other region.

For instance, the Physical Security Upgrade Programme (PSUP) only applies to GDNs in GB, whereas metering is an activity undertaken only by GDNs in NI and is not part of the regulated activities of GDNs in GB. Excluding these costs should improve the comparability of the data. The activities that have been excluded from the cost data are summarised below.

- Training & apprentice costs
- Streetworks (Traffic Management Act) costs
- Costs associated with the Physical Security Upgrade Programme (PSUP)
- Costs associated with gasholder decommissioning
- Environmental costs
- Land remediation costs
- Metering costs
- Network business rates
- Sales and marketing costs

A separate paper has been developed by the UR, which outlines the rationale of these exclusions in more detail.

Scale

The scale of the GB and NI GDNs is significantly different. For example, the average number of customers and network length is respectively 2.7m and 33,000km for GB GDNs whereas the respective numbers for NI GDNs[14] are 110,000 and 2,200km.

In the case of network utilities it is a widely held belief, supported by empirical evidence (e.g. RIIO-GD1), that economies of scale exist. However, the extent of such scale effects and at what level of scale they are present is an empirical issue. Further, it is not clear whether these scale effects can be characterised by a linear (or log-linear) functional form or more complex relationship. For example, an increase in a company's scale might have a large impact on unit costs when the scale is relatively small and a moderate impact when the company has reached a relatively large size. If the relationship between cost and scale is non-linear but the modelling approach imposes linearity then the coefficient estimates and efficiency scores may be biased.

In the context of the GB-NI benchmarking, there are a number of ways that differential scale effects could be controlled for or identified.

1. Assume that a log-linear (i.e. Cobb-Douglas) specification is a good approximation of the underlying relationship. However, as discussed above, if the approximation is invalid the efficiency estimates will be biased.

2. Include a simple quadratic term of the scale variable, which may be sufficient to capture these effects. However, given the small sample size, this approach might be unable to identify statistically significant effects.

3. Examine the sensitivity of the results across alternative sub-samples. For instance, if the estimated scale coefficient is sensitive to the exclusion of the NI GDNs from the estimation sample, then this would provide some indication of differential scale effects.

Network utilisation

The use of a CSV with fixed weights between volumes delivered, customer numbers and network length implicitly assumes that the ratio between these three factors does not materially affect costs or that the ratio is similar across all GDNs. Although this is virtually true for GB GDNs, NI GDNs have different network utilisation. This is particular the case for FE, which is quite idiosyncratic.

If the fixed CSV weights do not sufficiently capture the relationship between cost and the three scale variables, then the efficiency estimates might be biased. Models that deviate from the CSV approach have been used, whereby scale is controlled for by customer numbers, network length per customer and volume per customer.

[14] Average of PNGL and FE for 2015. No data for SGN Northern Ireland.

Network age and composition

PNGL was first granted a gas distribution license in 1996 whilst Firmus Energy first started operating in 2005. Both companies' networks are almost wholly constructed with modern polyethylene pipes. The networks that they operate serve the same purpose as their GB counterparts but the technical make up of their networks is materially different from GB GDNs. GB GDNs have been in operation much longer and have a significantly older stock of network assets, with large proportions still being made of iron which is more likely to leak than more modern polyethylene (PE) pipes.

Older network assets might have higher maintenance cost than a modern equivalent asset. If the different nature of GDN's network composition and asset age is not controlled for it may bias the estimation of operational cost efficiency and potentially overstate the efficiency of GDNs with more modern networks. In order to control for network age and composition, a variable that measures iron mains as a percentage of the total network is included in the model.

Estimation error

Econometric benchmarking assumes that all exogenous factors (i.e. outside management's control) that drive differences in cost performance are sufficiently accounted for within the model. If there are omitted factors or unobserved heterogeneity, then the model is mis-specified, which could induce bias in efficiency estimates. This is something that has been recognised in previous price controls and typically has been dealt with by:

1. Benchmarking comparators not to the efficiency frontier but to the upper decile or upper quartile of the efficiency distribution[15];
2. Conducting sensitivity analysis, for instance, looking at the sensitivity of the results to model specification; and
3. Averaging econometric estimates with estimates from bottom-up models.

This study looks at both (1) and (2). The UR in a separate paper provides estimates from a bottom-up model and envisages considering both econometric and bottom-up estimates in the efficiency determination.

[15] Ofgem and Ofwat typically use a benchmark defined by the upper quartile of the distribution.

3 Data

Data across a total of ten GDNs, eight GB and two NI, has been compiled by the UR. Annual data covering the period from 2009 to 2015 are available for GB GDNs. Data for the last five and three years have been used for PNGL and FE, respectively. The shorter sample period for the NI operators is driven by data availability and quality considerations.

The data covers a number of key variables:

- Adjusted OPEX based on costs reported by GDNs, in line with Regulatory Instructions and Guidelines;
- Total km of gas mains (network length);
- Number of customers;
- Volume of gas distributed;
- Regional wages; and
- Total km of iron gas mains.

All data have been provided by Ofgem (GB GDNs) or the NI GDNs. The only exception is the wage data, which have been obtained from the Office for National Statistics.

4 Empirical analysis

4.1 Coefficient estimates

This section presents the results of the econometric analyses including model coefficients and corresponding catch-up efficiency saving estimates. In total, thirty three models have been estimated in order to examine the sensitivity of the results to model specification, GDN sample and estimation method. In particular:

- Nine alternative model specifications have been considered using different explanatory variables and approaches to capture time and scale effects;
- Three estimation techniques have been applied: POLS (baseline), RE and SFA; and
- Three alternative samples have been used: (1) All GB and NI GDNs are included in the estimation; (2) GB and PNGL are included in the estimation, and (3) only GB GDNs are used. The motivation for the use of these different samples is to best understand potential differences in the underlying relationship, and in particular any differential scale effects between GB and NI GDNs.

The accuracy of individual models will be affected by data quality, the degree of comparability of GB and NI GDNs and the extent to which the models can control for GDN heterogeneity. Models including the NI GDNs in the sample could potentially benefit from the additional specific information on costs for smaller GDNs. However, there are questions on variations in the FE cost data and challenges associated with separating NI's, and in particular FE's, efficiency from heterogeneity, i.e. small scale, network utilisation.

Tables 1-3 present the coefficient estimates and t-statistics[16] across all the alternative specifications, estimation methods and samples. The dependent variable is OPEX deflated by a regional wage index. The key insights of the analysis are set out below.

- The coefficient of the CSV variable is always statistically significant with a coefficient typically ranging between 0.69 and 0.81. This suggests that a 1% increase in the scale of a GDN is expected to increase costs by 0.69%-0.81%.
- The only exception is Model 9, which includes a CSV square term and for which the scale effect cannot be directly interpreted from the CSV coefficient. Nonetheless, the square term is highly insignificant indicating that the relationship between cost and CSV could be sufficiently described by a log-linear function.
- Notwithstanding this, there is an indication that economies of scale might be more pronounced for NI GDNs. In models where either FE or both NI GDNs are excluded, the scale coefficient increases.
- There is little difference in the CSV coefficient estimates across the alternative estimation methods (Model 1 vs. Model 10 and 11).
- Models 7 and 8 capture the scale effects without using a CSV but rather number of customers and the ratio of network length per customer (Model 7) or ratio of volume per customer (Model 8). The advantage of these models is that they require no assumption

[16] Bootstrap standard errors instead of asymptotic standard errors have been used to account for the small sample. These are based on pairs-cluster bootstrap which resamples clusters with replacement and allows for intra-cluster correlation.

around the relative impact of the three scale variables (effectively the weights are estimated from the model). Network length per customer is always statistically insignificant whereas volume per customer is significant at the 10% level when both GB and NI GDNs are used in the estimation. However, this result should be interpreted with caution. Volume per customer exhibits little variation across GDNs and over time. It is mainly FE that has a considerably different volume per customer, therefore the impact of this variable might capture not only volume effects but also FE-specific effects, including relative efficiency.

- Time effects have been controlled for using either year dummies or a linear time trend. In models where CSV and time effects are only considered in the specification (Models 2 and 3), the time coefficients indicate that the GDNs within the sample have reduced costs over time, keeping everything else the same. When percentage of iron mains is included in the model, the magnitude and significance of time effects diminishes potentially reflecting collinearity issues, i.e. time effects may primarily capture the change in the GB network composition, and/or degrees of freedom and statistical power issues rather than time-specific effects or efficiency improvements across the sector.
- The OPEX has been normalised by a regional wage index, which effectively imposes a one-to-one restriction in the relationship between OPEX and wages. In order to test this restriction, the wage index has been included on the right-hand side of the equation (Model 6).[17] Given that the wage variable is statistically insignificant, the unit restriction cannot be rejected.
- The Bayesian Information Criterion[18] (BIC) is reported across each model and indicates that a model with a CSV and time trend (Model 3) captures the underlying relationships better than a CSV-only model or more complex models that allow for non-linear time effects and/or other factors, i.e. network composition.

[17] The dependent variable is $\ln C_{i,t} - \ln W_{i,t}$, where C is the actual cost and W the regional wage index. This effectively imposes a unit coefficient on the wage variable: $\ln C_{i,t} - \ln W_{i,t} = \alpha + \beta X_{i,t} + \varepsilon_{i,t}$ is equivalent to $\ln C_{i,t} = \alpha + \beta X_{i,t} + \gamma \ln W_{i,t} + \varepsilon_{i,t}, where\ \gamma = 1$. This restriction can be tested by adding $\ln W_{i,t}$ on the right-hand side of the equation. If the coefficient of wage is statistically insignificant then the one-to-one restriction cannot be rejected.

[18] BIC measures the model fit but penalises models that are over-parametrised and is considered more powerful than other model fit statistics like R-square, adjusted R-square and Akaike Information Criterion (AIC). The lower the BIC, the higher the model fit.

Table 1 : Model coefficient estimates; GB & NI GDNs sample

Dependent Variable: Ln(OPEX/Wage index) - Bootstrap standard errors

Model	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Estimation method	POLS	POLS	POLS	POLS	POLS	POLS	POLS	POLS	POLS	RE	SFA
Variables											
Ln(CSV)	0.754***	0.746***	0.745***	0.713***	0.713***	0.694***			0.437	0.738***	0.725***
	(17.28)	(16.79)	(16.59)	(9.40)	(9.62)	(7.33)			(0.05)	(13.71)	(33.80)
Year dummy - 2009		0.157***		0.123*							
		(2.68)		(1.78)							
Year dummy - 2010		0.131**		0.107*							
		(2.29)		(1.77)							
Year dummy - 2011		0.0878		0.0698							
		(1.60)		(1.20)							
Year dummy - 2012		0.104**		0.0909*							
		(2.17)		(1.84)							
Year dummy - 2013		0.0808**		0.0722*							
		(2.21)		(1.95)							
Year dummy - 2014		0.0329		0.0283							
		(1.12)		(0.93)							
Time trend			-0.0241**		-0.0189					-0.0251**	-0.0261***
			(-2.39)		(-1.63)					(-2.42)	(-5.13)
Iron mains, %				0.404	0.400	0.822	1.053*	0.313			
				(0.69)	(0.66)	(1.29)	(1.71)	(0.49)			
Ln(Wage index)						-0.217					
						(-0.34)					
Ln(customer numbers)							0.703***	0.703***			
							(6.96)	(8.32)			
Ln(network length/customer numbers)							0.442				
							(1.16)				
Ln(volume/customer numbers)								0.319*			
								(1.91)			
Ln(CSV^2)									0.0123		
									(0.04)		
Constant	-12.62***	-12.59***	36.06*	-12.22***	25.95	-10.66***	-10.45***	-13.23***	-10.62	38.20*	40.11***
	(-20.12)	(-19.83)	(1.77)	(-11.82)	(1.11)	(-3.07)	(-7.82)	(-9.23)	(-0.15)	(1.81)	(3.92)
R-square	0.981	0.984	0.983	0.984	0.984	0.983	0.982	0.983	0.981	0.983	.
BIC	-79.643	-65.854	-85.702	-64.861	-84.611	-81.079	-75.941	-78.601	-76.560	.	.
N	65	65	65	65	65	65	65	65	65	65	65
T	10	10	10	10	10	10	10	10	10	10	10

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Model coefficient estimates; GB GDNs & PNGL sample

Dependent Variable: Ln(OPEX/Wage index) - Bootstrap standard errors

Model	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Estimation method	POLS	POLS	POLS	POLS	POLS	POLS	POLS	POLS	POLS	RE	SFA
Variables											
Ln(CSV)	0.769***	0.765***	0.764***	0.733***	0.733***	0.703***			0.723	0.759***	0.760***
	(9.44)	(9.30)	(9.43)	(6.43)	(6.30)	(5.17)			(0.03)	(7.74)	(22.16)
Year dummy - 2009		0.166***		0.138							
		(2.80)		(1.63)							
Year dummy - 2010		0.146**		0.125*							
		(2.53)		(1.72)							
Year dummy - 2011		0.103*		0.0866							
		(1.87)		(1.30)							
Year dummy - 2012		0.120**		0.107*							
		(2.49)		(1.96)							
Year dummy - 2013		0.0784*		0.0714							
		(1.94)		(1.61)							
Year dummy - 2014		0.0586***		0.0547***							
		(3.20)		(2.76)							
Time trend			-0.0250**		-0.0205					-0.0250**	-0.0250***
			(-2.42)		(-1.48)					(-2.45)	(-5.13)
Iron mains, %				0.324	0.321	0.784	0.827	0.311			
				(0.41)	(0.42)	(1.08)	(1.13)	(0.48)			
Ln(Wage index)						-0.209					
						(-0.29)					
Ln(customer numbers)							0.715***	0.704***			
							(5.89)	(5.79)			
Ln(network length/customer numbers)							0.323				
							(0.63)				
Ln(volume/customer numbers)								0.315			
								(1.48)			
Ln(CSV^2)									0.00172		
									(0.00)		
Constant	-12.83***	-12.87***	37.53*	-12.50***	28.86	-10.83***	-11.08***	-13.22***	-12.53	37.67*	37.46***
	(-10.91)	(-10.81)	(1.80)	(-7.71)	(1.03)	(-2.83)	(-4.70)	(-7.00)	(-0.08)	(1.82)	(3.80)
R-square	0.966	0.972	0.971	0.973	0.972	0.970	0.968	0.969	0.966	0.971	.
BIC	-77.014	-64.541	-84.026	-62.340	-81.760	-77.312	-73.189	-74.749	-72.893	.	.
N	62	62	62	62	62	62	62	62	62	62	62
T	9	9	9	9	9	9	9	9	9	9	9

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Model coefficient estimates; GB GDNs sample

Dependent Variable: Ln(OPEX/Wage index) - Bootstrap standard errors

Model	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Estimation method	POLS	POLS	POLS	POLS	POLS	POLS	POLS	POLS	POLS	RE	SFA
Variables											
Ln(CSV)	0.775***	0.769***	0.767***	0.788***	0.787***	0.805***			11.58	0.715***	0.754***
	(5.08)	(5.23)	(5.07)	(4.21)	(3.98)	(4.21)			(0.28)	(4.13)	(5.53)
Year dummy - 2009		0.169***		0.120							
		(2.62)		(1.02)							
Year dummy - 2010		0.146**		0.104							
		(2.22)		(1.03)							
Year dummy - 2011		0.109*		0.0759							
		(1.75)		(0.86)							
Year dummy - 2012		0.127**		0.103							
		(2.32)		(1.50)							
Year dummy - 2013		0.0766*		0.0617							
		(1.67)		(1.13)							
Year dummy - 2014		0.0676***		0.0602***							
		(3.68)		(2.83)							
Time trend			-0.0249**		-0.0166					-0.0252**	-0.0250***
			(-2.30)		(-0.86)					(-2.27)	(-4.67)
Iron mains, %				0.558	0.552	1.150	0.920	0.596			
				(0.43)	(0.44)	(1.60)	(1.16)	(0.84)			
Ln(Wage index)						-0.253					
						(-0.33)					
Ln(customer numbers)							0.785***	0.810***			
							(3.84)	(4.82)			
Ln(network length/customer numbers)							0.248				
							(0.32)				
Ln(volume/customer numbers)								0.358			
								(1.58)			
Ln(CSV^2)									-0.371		
									(-0.26)		
Constant	-12.93***	-12.93***	37.25*	-13.36***	20.07	-12.16***	-12.47***	-15.03***	-91.52	38.70*	37.45***
	(-5.86)	(-6.05)	(1.72)	(-4.75)	(0.50)	(-2.91)	(-3.23)	(-6.21)	(-0.30)	(1.73)	(3.34)
R-square	0.749	0.792	0.788	0.804	0.799	0.798	0.776	0.790	0.759	0.788	.
BIC	-64.405	-50.785	-69.742	-49.946	-68.767	-68.427	-62.696	-66.334	-62.496	.	.
N	56	56	56	56	56	56	56	56	56	56	56
T	8	8	8	8	8	8	8	8	8	8	8

t statistics in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.2 Catch-up efficiency

Table 4 sets out the catch-up efficiency savings. These are computed using the third most efficient GDN as the benchmark.[19]

Two sets of results are presented. The 2014 catch-up efficiency is computed using the 2014 model residuals, whereas the 2015 values are calculated using the 2015 residuals. The 2015 NI cost data are not actual data but estimates provided by the NI GDNs. Furthermore, it is useful to examine the stability of the results across different years. Although in principle using the latest available year is more appropriate from an economic point of view, data quality and statistical error may lead to significant over/under-estimation of the catch-up efficiency.[20]

Relative efficiency values are presented for the best four models.[21]

- **Model 3.** From a statistical point of view, Model 3 provides the best fit in terms of BIC across all samples. Notwithstanding this, differences in the fit of different models might not be statistically significant, whereas other models could still contain useful information and therefore should still be considered.[22]
- **Model 2.** Although the time dummy specification is rejected by the data on the basis of parsimony, the time dummy coefficients and statistical significance indicate non-linear time effects.
- **Model 5.** This model suggests that network composition is statistically insignificant but this might be associated with the small sample and the model's inability to identify significant effects. This model is considered to entertain the possibility that network composition is economically significant.
- **Model 10.** This is similar to Model 3 but estimated using RE.

The key findings are summarised below.

- The results, in particular for FE are quite sensitive to the year used as the basis for the computation of efficiency. The underlying volatility is likely to reflect data issues rather than actual change in efficiency.
- The results are also sensitive to the sample used for the estimation and whether the age or state of networks is accounted for, or not. The 2014 FE efficiency ranges from 5.1% to 8.7% when both GB and PNGL data are used in the estimation and between 0% and 29% when GB only data are considered. For PNGL, the range of efficiencies is 5.8%-8.1% and 2.6%-25% in the GB and PNGL, and GB only samples, respectively. However, the results are noticeably more stable for Model 3.
- The highest efficiency opportunities for both FE and PNGL are documented in Model 5, which includes the network composition variable. If this variable is significant but omitted from the model, then relative efficiency will be under-estimated for the NI GDNs.

Coefficient and, in particular 2015 efficiency estimates are expected to be revised in the final determination where actual 2015 costs, rather than estimated, will be used.

[19] This is equivalent to the upper quartile used in previous studies and aims to mitigate potential estimation error as discussed in section 2.6.

[20] SFA estimates assume time-invariant inefficiency, therefore the corresponding values reflect average sample, rather than 2015, efficiency. More sophisticated SFA models which allow for time variation in efficiency are available, however these are more complex and typically are used when the sample size is considerably larger than that available in this study.

[21] Evaluation of alternative models should also consider model diagnostics. Diagnostic tests will be considered in the final determination.

[22] There are no standard tests to evaluate the statistical difference of BIC.

Table 4: Catch-up efficiency estimates (relative to the third most efficient GDN)

	FE			
Model	**2014**		**2015**	
	GB & PNGL sample	GB sample	GB & PNGL sample	GB sample
2	0.087	0.102	0.268	0.280
3	0.085	0.097	0.267	0.276
5	0.051	0.292	0.238	0.426
10	0.066	0.000	0.251	0.117
	PNGL			
Model	**2014**		**2015**	
	GB & PNGL sample	GB sample	GB & PNGL sample	GB sample
2	0.071	0.081	0.107	0.117
3	0.070	0.078	0.106	0.114
5	0.081	0.255	0.114	0.277
10	0.058	0.026	0.095	0.000

4.3 CSV sensitivity analysis

A CSV with weights equal to 0.5, 0.25 and 0.25 for customer numbers, network length and volume has been used as baseline, based on UR's view of the relative importance of the different scale measures in terms of operating costs.

In this section, the suitability of those weights is examined from a statistical view point. A series of alternative models have been estimated using all possible combinations of weights (using increments of 0.05) and comparing the fit of these models as measured by R-square. The results of this exercise are shown in Figure 3.

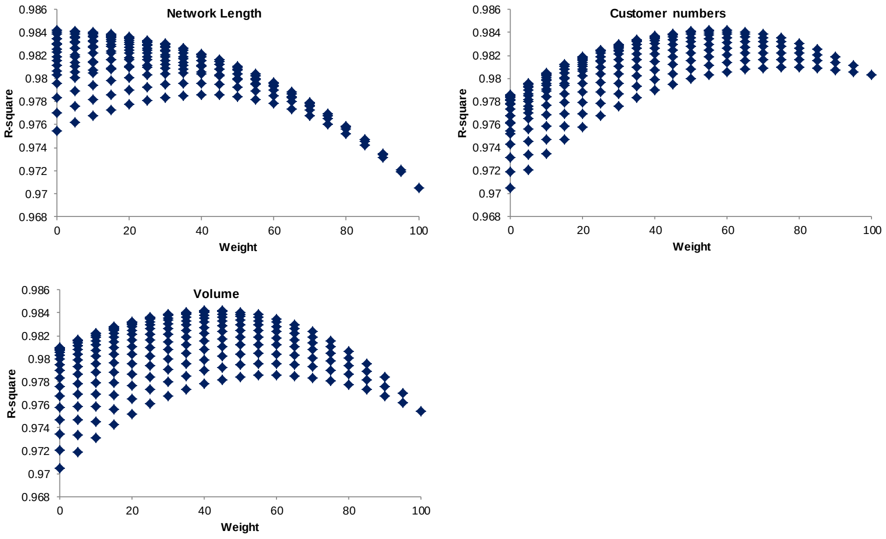
For each specific weight assigned to a scale variable (0.50), there are multiple combinations of weights for the other two variables (0.25/0.25 or 0.1/0.4), hence the multiple R-square values per weight reported.

This analysis provides two key insights:

- First, the choice of weights has little impact on model fit. R-square ranges between 97% and 98.5% and is virtually the same across the models; and
- The baseline weights deliver a fit close to the top of the distribution with an R-square equal to 98.2%.

However, these results should be interpreted with care as they mainly reflect the high correlation of the underlying variables.

Figure 3: CSV weights and model fit



Finally, the sensitivity of relative efficiency was examined across a number of alternative weights which placed more weight on network length and volume. These models suggest that the efficiency estimates are sensitive to the choice of weights, which was expected given the difference in the network utilisation between PNGL, FE and their GB counterparts. The UR is of the view that customer numbers is a more important cost driver than network length and volume, and should have greater weight as reflected in the baseline weights. This is also consistent with precedence and other independent analysis. Ofgem in RIIO-GD1 used workload measures related to specific services (repex, connections, emergency services) which are mainly a function of number of customers. Further, National Grid has shown that OPEX is mainly driven by number of customers rather than network length.[23] Notwithstanding this, given the idiosyncrasy of the NI GDNs and in particular FE, and the sensitivity of the results, special factors (both positive and negative) should be taken into account for the efficiency determination.

[23] https://www.ofgem.gov.uk/sites/default/files/docs/2007/07/ngg---response---section-3---non-confidential.pdf.

5 Appendix

5.1 Efficiency computation

Let the actual costs of company i in period t be described by the following process:

$$C_{it} = f(x_{it}, \theta)\exp(\epsilon_{it}) \tag{1}$$

where x_{it}, are the inputs, θ are the parameters and ϵ_{it} is a mean-zero error-term which represents deviations in cost efficiency. Let ϵ_{pt} denote the p^{th} percentile of the error-term distribution, then the cost-efficiency of company i relative to efficiency at x_{it} and ϵ_{pt} is:

$$\phi_{it}^{1-p} = \frac{C_{pt}}{C_{it}} = \exp[-(\varepsilon_{it} - \epsilon_p)] \tag{2}$$

where ϕ_i^{1-p} denotes the efficiency relative to the $1 - p^{th}$ percentile of the *efficiency distribution*. For example, if $\phi_i^{0.9} = 85\%$, then the producer could save 15% of their costs if they became as efficient at the lowest 10th percentile of the error-distribution. Equation (2) is implemented by replacing the parameters with consistent estimates: $\hat{\varepsilon}_{it} = \log C_{it}/f(x_{it}, \hat{\theta})$ and restricting the potential maximum to 100%. Thus the relative cost savings are:

$$1 - \hat{\phi}_{it}^{1-p} = 1 - \min\{1, \exp[-(\hat{\varepsilon}_{it} - \hat{\varepsilon}_{pt})]\} \tag{3}$$