Regional factors for RIIO-GD3: Regional wages
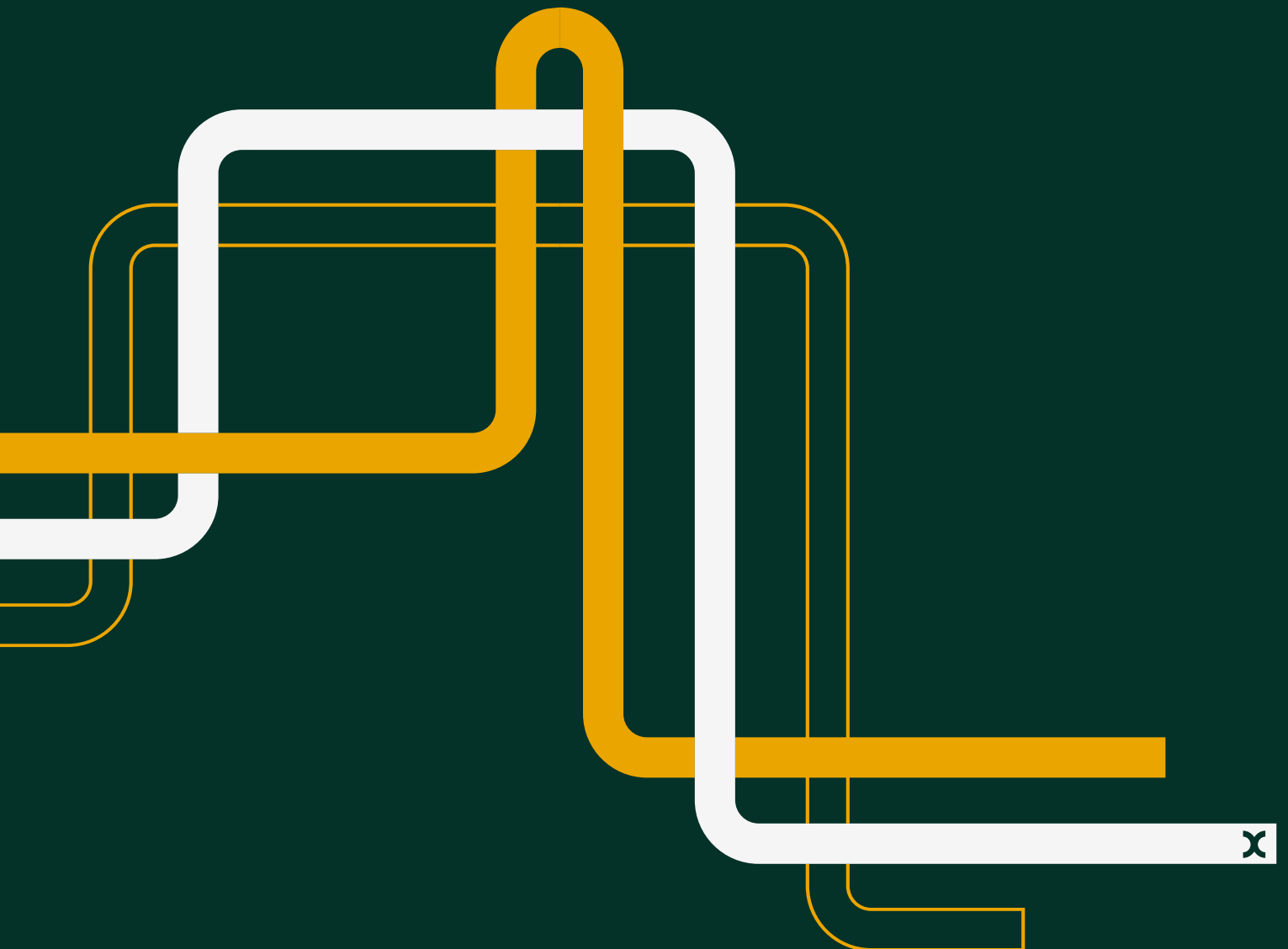
—

Prepared for Wales & West Utilities

22 November 2024

oxera

Contents

Executive summary

Ofgem's RIIO-3 Sector Specific Methodology Decision for gas distribution (GD) indicates plans to re-evaluate the regulator's regional wage adjustment methodology for gas distribution networks (GDNs). While Ofgem rightly highlights wage convergence between the London region and the rest of the UK,[1] we find that its current two-digit Standard Occupational Classification (SOC) methodology results in inaccurate adjustments. In particular, due to its level of aggregation, Ofgem's current approach results in the inclusion of 71 three-digit SOC categories, of which only 39 are relevant to GDN activities, resulting in distorted regional wage adjustments.

A more granular approach, using three-digit SOC codes, would better capture actual GDN occupations. Ofgem moved to two-digit SOC codes, rather than the three-digit ones used at GD1, in order to 'reduce uncertainty and missing data in the ASHE wage estimates' [sic].[2] However, we find that concerns about missing data are negligible (fewer than 0.7% of the relevant observations). For the limited relevant cases, we propose a simple method to estimate any missing three-digit-level information. However, given that there is such a negligible amount of missing data, an index based on three-digit SOC codes offers the most accurate adjustment (irrespective of the preferred treatment of missing values).[3]

Forecasting methods also warrant a re-examination and possible revision. In GD2, Ofgem averaged wage indices from 2015–19 to forecast adjustments over GD2. While this approach may have seemed appropriate at the time, GDNs London, the East of England and Southern's wage indices have declined since COVID-19 (as is evident in the figure below). This suggests that alternative approaches should be considered for GD3, such as rolling forward the latest data point or using a shorter three-year historical average (if not trends). The question of the most appropriate approach can be revisited when 2024 ONS data becomes available, in order to confirm trends.

Failing to revise the forecast approach and methodology risks significantly overcompensating some GDNs for the regional wage

---

[1] Ofgem (2024), 'RIIO-3 Sector Specific Methodology Decision – GD Annex', 18 July, paras 5.46–5.47.
[2] Ofgem (2020), 'RIIO-GD2 Final Determinations: Step-by-Step Guide to Cost Assessment', 8 December, Appendix A, Table 7.
[3] When rounding to two decimal points, we found that it makes no difference whether missing data points are ignored or inferred.

adjustment required. For instance, updating indices based on Ofgem's current two-digit methodology with updated data to 2023 and rolling forward the latest three-year average over 2024–26, we estimate that Cadent's London GDN received a £19.5m (or 24%) higher pre-modelling adjustment than warranted ex post. Under our proposed three-digit approach (and similar three-year average forecast), this estimated overcompensation rises to £28.5m (or 39%).

Moreover, moving to a three-digit approach would have mitigated this overcompensation/forecast error to some extent. For example, if Ofgem's ex ante forecasts had been instead based on the equivalent three-digit index at the time (even if still using a five-year average), this would have reduced the estimated overcompensation above for London to only 6% or 19% (rather than 24% or 39%) based on the respective two- and three-digit outturn.

The GD2 experience underscores the need for improved forecasting and methodology to ensure fairer, more accurate adjustments, better comparability, and a balanced risk–reward framework across GDNs.

Updated and alternative wage indices for London GDNs
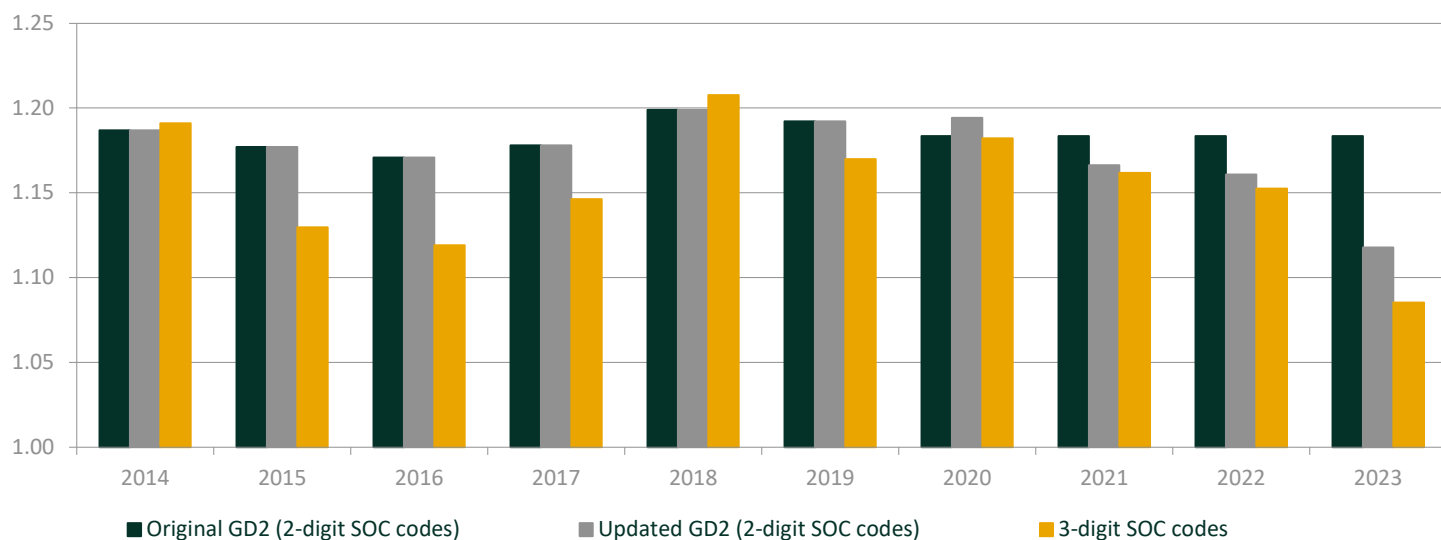
Note: ONS data for 2023 is provisional.
Source: Oxera based on ONS ASHE tables 3a, 15.5a and 14.5a.

1   Introduction

In the cost assessment for RIIO-GD3, Ofgem will need to benchmark the costs of the gas distribution networks (GDNs) against each other. In order to ensure a like-for-like comparison, any regional- or company-specific factors that result in material cost differences across the GDNs should be captured, either within the cost drivers in the model or using pre-modelling adjustments.

Ofgem's recent RIIO-3 Sector Specific Methodology Decision (SSMD) for gas distribution (GD) summarises the regulator's intended modelling approach for GD3. While Ofgem has not made any definitive decisions, it has indicated that the GD2 framework will largely form the basis of the GD3 approach and the specific model testing and alterations that it is considering. For regional factors, Ofgem highlights the following two specific considerations and potential changes for GD3.

- **Regional wages recalibration**, revaluating the activities covered and noting that London wages have increased more slowly than those in the rest of the UK.
- **Testing within-model density controls**, testing a density (and density squared) variable within model to account for both urbanity and sparsity effects.[4]

We agree that regional wage convergence and sparsity/urbanity effects are the two main categories of regional factors that should be re-examined and potentially changed, as we consider that both are unlikely to be either sufficiently or correctly captured in Ofgem's existing cost drivers.[5] This report focuses on regional wages, with sparsity effects discussed in the accompanying report (Oxera, 2024).[6]

While we note that Ofgem has correctly noted the wage convergence trends, we find that the current two-digit-based methodology is inappropriate because it includes irrelevant three-digit labour categories. We discuss why this is the case, and propose remedies to account for both this and wage convergence, below.

---

[4] Ofgem (2024), 'RIIO-3 Sector Specific Methodology Decision – GD Annex', 18 July, paras 5.46–5.47.
[5] As the final GD3 models are not yet known, this would need to be revaluated when Ofgem has published these models.
[6] Oxera (2024), 'Regional factors for RIIO-GD3: Sparsity', November, Report prepared for Wales & West Utilities.

2   Context

At GD2, Ofgem calculated the regional wage index based on regional occupational data using two-digit Standard Occupational Classification (SOC) codes. These codes are weighted using industry-average occupational weights based on the number of full-time equivalent employees (FTEs). Ofgem moved to two-digit SOC codes, rather than the three-digit ones used at GD1, in order to 'reduce uncertainty and missing data in the ASHE wage estimates' [sic].[7]

However, two-digit SOC codes are fairly broad categories that are unlikely to accurately reflect the occupations of people working at the GDNs. For instance, SOC 31, 'Science, engineering and technology associate professionals', is split into three further sub-categories: 'Science, engineering and production technicians' (311), 'Draughtspersons and related architectural technicians' (312) and 'Information technology technicians' (313). These three sub-categories have different wage levels[8] and different regional variations.[9] Moreover, the sub-category 'Draughtspersons and related architectural technicians' is not relevant to the GDNs.

The two-digit SOC code methodology that Ofgem applies therefore picks up regional wage differences for sub-sectors that are not particularly relevant to the GDNs. Overall, Ofgem considers 39 of the ONS's three-digit codes to be relevant to GDNs. However, when using the two-digit methodology, this results in 71 three-digit occupations feeding into the calculations. Therefore, **almost half of the three-digit occupations that are implicitly being used in Ofgem's two-digit method are not relevant to GDNs**. This means that the resulting wage index applied by Ofgem produces incorrect adjustments that do not reflect the actual wage pressures experienced by GDNs.

At GD2 final determinations (FD), Ofgem had ONS data up to 2019. All future indices were set equal to the average of the last five years (2015–19). As historical years were quite similar and there was no clear trend, this method was not controversial. As discussed in section 5, this may also no longer be the appropriate forecasting approach at GD3.

---

[7] Ofgem (2020), 'RIIO-GD2 Final Determinations: Step-by-Step Guide to Cost Assessment', 8 December, Appendix A, Table 7.
[8] At a UK level, wages for the first category are lower than those for the other two.
[9] For example, for the East in 2022, wage levels for occupation 311 are higher than the UK-wide average figure, but for 312 they are lower than the UK figure.

3   Suggested (three-digit) methodology

Using more granular SOC codes has the advantage of better reflecting the actual occupations undertaken at the GDNs. A possible disadvantage, at least in theory, is that the more granular the SOC codes are, the greater the likelihood is that there is missing data, because the ONS does not have sufficiently robust granular wage information.

However, in practice, we find that the three-digit SOC codes have very little missing wage information for the occupations that receive high occupational weightings when constructing the regional wage indices for the GDNs' activities. That is, only 1.2% of the 3,432 relevant observations are missing, and less than 0.7% are missing on an industry-weighted average basis (given the small weighting of the relevant observations—predominantly in the 'Chief Executives and Senior Officials' category[10]).

Where this information is missing, a simple method can be applied to estimate any missing three-digit-level information.[11] To infer these missing values, we do the following:

- identify relevant missing values (i.e. for three-digit occupations that have a positive weighting);
- calculate the difference between the three-digit and the corresponding two-digit wage at a national level;
- apply the national-level wage differential to the two-digit wage of the region where the three-digit wage was missing. This allows us to infer the three-digit wage for this area.

A stylised example of this method is as follows. The two-digit wage for a particular region is £100 but the three-digit wage data is missing. At the UK level, we have two-digit and three-digit wages for the corresponding SOC codes. The two-digit UK wage is £200 and the three-digit UK wage is £180. Our methodology calculates the missing three-digit wage for this region as 100 x (180/200) = £90. That is, we know from the UK-level data that the three-digit wage is likely to be somewhat lower than the

---

[10] Of the 40 missing observations between 2014 and 2021, across 11 regions, 27 are for the Chief Executives and Senior Officials category. This category has a weight of only c. 0.2% for any given year–region pair. The other 13 missing observations are similarly spread across six other three-digit-level occupations, which generally also have very small weightings.
[11] For details on these calculations, an accompanying Excel spreadsheet has been provided to Ofgem.

two-digit wage. The national percentage difference, as a proxy, is thus applied to the two-digit regional wage to infer the missing value.

Moreover, given that there is such limited relevant missing data (less than 1%), we find that this methodology makes little difference compared with simply ignoring these missing data points.[12]

Therefore, irrespective of the preferred treatment of missing values, we find that using the three-digit-level data should be considered by Ofgem as the most accurate way of calculating this adjustment.

---

[12] When rounding to two decimal points, we found that it makes no difference whether missing data points are ignored or inferred.

4 Results

Figure 4.1 below shows the wage index for Cadent's London GDN, as the GDN with the highest regional wage index, over time. The index is calibrated to show the wage levels relative to one, such that GDNs that do not serve customers in London or the South East of England all have an index of one (with only London—Lon, East of England—EoE, and Southern—So having indices greater than one). Figure 4.1 compares three indices for Lon, specifically, as follows.

- **Original GD2 (two-digit)**: the regional wage index estimated by Ofgem at the time of the GD2 FD—which forecast values over 2020–26 based on the five-year average of the latest outturn data at the time (2015–19).
- **Updated GD2 (two-digit)**: updating the GD2 two-digit index with the latest data from the ONS ASHE data up to 2023.[13]
- **Three-digit approach**: showing the index values resulting from our suggested approach above and the 2023 updated ONS data.

Figure 4.1 clearly shows wage convergence across the industry: when updating Ofgem's two-digit methodology, the difference between wages in London and in the rest of the country has decreased in recent years (compared with Ofgem's previous forecasts at GD2). Ofgem has noted the same trend from its own regional factor updates during the GD3 cost assessment working groups (CAWGs) process.[14]

Using the three-digit SOC codes further reduces the wage premium for GDNs such as Lon, albeit this has a small impact compared with simply updating the existing methodology.

---

[13] That is, the ASHE series over 2020–23, where 2020–22 are revised editions and 2023 is the provisional data (the latest version for which three-digit data is available at the time of writing).

[14] Ofgem (2024), 'RIIO-GD3 Cost Assessment Working Group 7. TOTEX modelling and BPDT development', 10 April, slide 19. Due to the change in the ONS ASHE series from SOC10 to SOC20 codes from 2020, Ofgem has remapped certain four-digit codes for its provisional update of the wage series, to reconstruct comparable wage indices over time. As this remapping is still provisional, we do not attempt to replicate it here. Instead, we use the two-digit and three-digit results directly from the ONS ASHE series. This has no impact on our conclusions, as the results from Ofgem's analysis with the remapping and our analysis without it are materially the same (that is, there is at most a 0.01 index difference for a select few GDN-year instances, and the resulting trends are exactly the same).

Figure 4.1   Updated and alternative regional wage indices for Lon GDNs
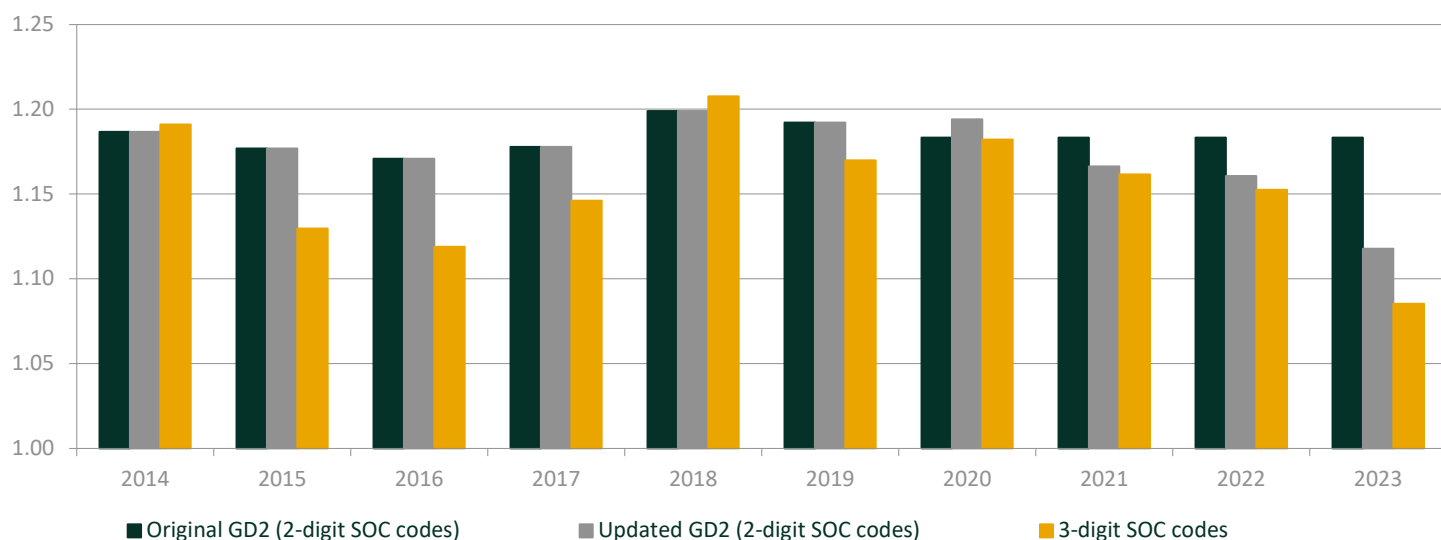
Note: ONS data for 2023 is provisional.
Source: Oxera based on ONS ASHE tables 3a, 15.5a and 14.5a.

One reason for the downward trend in the relative wage differential for the London region is likely to be the more flexible work environment and increase in remote working since COVID-19. This reduces the need for employees to be physically located in high-cost areas (e.g. London). For example, recent ONS research shows that 28% of UK adults have hybrid working arrangements (and c. 12% work fully remotely)—a trend that has remained fairly stable post pandemic.[15]

This is consistent with the operational trends that WWU has been observing. For example, we understand that WWU has faced increased competition from other GDNs for REPEX contractors who used to work only within their region. We understand from WWU wage-setting and bargaining dynamics have also changed across the sector, and that many organisations in the sector no longer implement regional weightings in light of the increase in hybrid and flexible working (especially for back office functions).

In addition to the choice of index, forecasting methods warrant a re-examination and possible revision. As is evident from Figure 4.1, there is now a clear downward trend in the wage index for London. While a linear extrapolation of this trend might be too extreme,[16] Ofgem could apply the latest data point (here, 2023) to all forecast years, or use a shorter-

---

[15] Office for National Statistics (2024), 'Who are the hybrid workers', 11 November.
[16] This should be confirmed once the full ONS ASHE data for 2024 is available.

period average to account for the lower relative wage indices since the COVID-19 pandemic (e.g. the three years 2021–23). The former would be consistent with Ofwat's approach to some of the external population characteristic indices used in its modelling (e.g. for the deprivation metrics used in its retail services cost benchmarking).

To see the importance of choosing the appropriate forecast method, note that Lon, So and EoE have each received much greater pre-modelling adjustments at GD2 FDs than are justified by the outturn environment since (i.e. there was a significant forecast error). Table 4.1 below shows the labour adjustments for Lon, So and EoE under each of the approaches discussed above, from left to right, as follows.

- **Original GD2 (two-digit)**: Ofgem's FD approach, using data up to 2019 and forecasts based on the five-year preceding average.
- **Updated two-digit (last year forecast)**: updating the GD2 index up to 2023 and rolling forward the last year as forecast.
- **Updated two-digit (three-year forecast)**: as above, but with 2021–23 average as forecast over 2024–26.
- **Alternative GD2 (three-digit)**: our suggested approach, as at 2019 (with forecasts based on the five-year preceding average).
- **Three-digit (last year forecast)**: our suggested approach up to 2023 and rolling forward the last year as forecast.
- **Three-digit (three-year forecast)**: as above, but with 2021–23 average as forecast over 2024–26.

The extent of the forecast error can be seen in comparing the original GD2 index-implied adjustments with those of the updated indices. For example, based on updating Ofgem's two-digit methodology up to 2023 and rolling forward the three-year average, Lon received a c. £19.5m (or 24%) greater pre-modelling adjustment than is justified ex post.[17] Similarly, based on what we propose to be the more accurate three-digit alternative approach,[18] Lon received a c. £28.5m (or 39%) greater pre-modelling adjustment than the outturn environment has justified.

The three-digit methodology proposed would have mitigated this forecast error to some extent. If Ofgem's ex ante forecasts were instead based on the equivalent three-digit index at the time (even if still a five-year average), this would have reduced the forecast error for Lon

---

[17] That is, from Table 4.1, the £101.63m - £82.06m = £19.57m (or c. 24% of £82.06m).
[18] That is, also updated and rolled forward based on the three-year average. From Table 4.1, the £101.63m - £73.08m = £28.55m (or c. 39% of £73.08m).

to only 6% or 19% (not 24% and 39%) more than the respective two- and three-digit outturn (and three-year average forecasts) suggest.

GDN	Two-digit			Three-digit		
	Original GD2	Updated, last year forecast	Updated, three-year forecast	Alternative GD2	Updated, last year forecast	Updated, three-year forecast
EoE	9.86	6.72	7.68	8.26	5.32	6.83
Lon	101.63	71.77	82.06	86.92	56.49	73.08
So	88.99	58.93	66.68	81.04	38.16	50.19

Source: Oxera based on Ofgem's updated cost assessment dataset (November 2023).

We note that the cost modelling benchmark (whether upper quartile or 85th percentile) should not be detrimentally affected by the change in regional wage indices, given that the regional wage adjustments affect only GDNs that have historically not performed near the benchmark level.[19]

---

[19] The benchmark would be expected to be affected only if the GDNs ranked between second and third had lower regional wage adjustments, and thus higher costs fed into the regression models. Based on Ofgem's latest modelling based on outturn data up to 2023, EoE, Lon and So rank seventh, eight and sixth respectively. Ofgem (2024), 'RIIO-GD3 Cost Assessment Working Group 7. Totex modelling and BPDT development', 10 April, slide 11.

5   Implications for Ofgem's approach

First and foremost, we find that a three-digit-level approach would provide a more accurate basis for a regional cost adjustment than Ofgem's two-digit approach at GD2. There is very limited relevant missing data, and where this is the case a simple method can be used to infer missing values. This is preferable to a counterfactual of including a large number of occupations that are not relevant to GDN operations.

The GD2 experience has also shown that an incorrect forecast period risks over- or undercompensating the regional wage adjustment required by some GDNs. This is to the detriment of Ofgem's modelling framework, affecting both its ability to compare on a like-for-like basis and the overall balance of risk and reward across GDNs.

The appropriate forecast approach can be reconsidered when the 2024 ONS ASHE data is available, in order to assess whether the downward trend has stabilised or continued post COVID-19. This, in turn, would inform whether a linear trend or rolling forward of a (shorter) historical average or latest annual value is more appropriate.

oxera

Contact

Alan Horncastle
Partner
+44 (0) 1865 253015
alan.horncastle@oxera.com

oxera.com