13 May 2024

Economic Insight Ltd

ONGOING EFFICIENCY FOR GAS NETWORKS AT RIIO-3

A report for the Energy Networks Association

CONTENTS

1 Introduction and executive summary

In this report, we undertake a benchmarking analysis using total factor productivity (TFP) data to arrive at estimates for an appropriate ongoing efficiency (OE) challenge for gas networks at RIIO-3. Having developed a robust and evidence-based approach, we conclude that OE at RIIO-3 for gas networks will most plausibly be in the range of 0.2% to 0.8% (the midpoint being 0.5%). We do not consider it appropriate to make post-benchmarking adjustments to the range, or to select values at the extreme ends of the range.

This report should be read in conjunction with our October 2024 supplementary report[1], which provides further evidence on OE for RIIO-3.

1A. The existing approach to determining OE reveals the challenges in its estimation and should be reconsidered

The overall approach to estimating OE is well established

OE is the efficiency (or productivity) improvement that even the most efficient company in an industry can achieve. As such, all companies have scope to make OE gains over time.

Accordingly, as part of its price setting framework, Ofgem will set an OE challenge at RIIO-3, which will apply to all transmission and gas distribution companies. At a high level, the existing method used by Ofgem (and other sectoral regulators) for setting OE is relatively well established, and consists of two main elements:

- Building a range of OE estimates from benchmarking of productivity data (EU KLEMS being a commonly used source).

[1] *'Further Evidence on OE for Gas Networks at RIIO-3.' Economic Insight (October 2024).*

- Selecting a point from within the estimated range, sometimes with further post-benchmarking adjustments (i.e. adjustments applied to the OE estimate derived from the benchmarking exercise in the above bullet).

Applying the method at a detailed level is inherently challenging

Whilst the above two methodological steps may seem straightforward, in practice, estimating the *'correct'* OE target in regulated industries remains inherently challenging, because:

- **OE cannot be directly observed within the available data used for benchmarking, which provides measures of productivity.** Productivity growth can be driven by many different components of productivity, not just OE (for example, catch-up efficiency and economies of scale). This raises questions as to how any estimate derived from benchmarking should be interpreted, for the purpose of setting an OE challenge.

- **The measurement of productivity in the first place is itself challenging and the data sources used are characterised by a degree of volatility, particularly at an individual industry level.** Notwithstanding the first issue above, the available productivity data reveals the intrinsic challenges in productivity estimation. At their core, productivity metrics measure changes in outputs, relative to changes in inputs. To create these metrics, one must therefore accurately measure the value of inputs and outputs. In addition, one must consider that those inputs and outputs may be volatile (especially at an industry level). For example, an industry with *'lumpy'* capex, and where investment leads to steady state increases in output in the long run, might exhibit:

  - a material *reduction* in productivity in the data at the time at which investment is made;

  - followed by a *large increase* in productivity, once output consequently increases;

  - then followed by *flat productivity*, if the output increase is sustained.

  Importantly in the above example, the underlying productivity of the industry is unchanged.[2]

- **A large number of complex analytical choices are required under any benchmarking approach**. Specifically, one must make choices as to: which measure of productivity to use (gross output, GO, or value added, VA); which comparator industries provide the most appropriate benchmarks; and the time period over which any benchmarking should be undertaken.

---

[2] *Whilst this example of volatility might be mitigated by using longer estimation windows, the underlying data shows that large up or downswings in inputs / outputs can persist for multiple years.*

- **Selecting a point estimate of OE from within the final range, and / or making post-benchmarking adjustments, requires a judgement call.** This risks adding an element of subjectivity around regulatory OE estimations. In turn, this can lead to perceptions of '*cherry picking*', if not properly evidenced.

These challenges are evident in regulator-determined OE, which is at odds with observable data and intuition

The challenges described above are evident in the regulatory history of setting OE.

Firstly, we observe that decisions made by regulators, based on benchmarking, are increasingly out of alignment with the reality of productivity growth in the UK. Specifically, it is readily observable that the UK has experienced a prolonged period of flat (and near-zero) productivity growth since 2008. Moreover, this is observable: (i) across most sectors of the UK economy; and (ii) more widely for many Western economies. This has been the topic of considerable discussion (widely labelled the *'productivity puzzle'*). In contrast:

- the average OE challenge set by sectoral regulators in the UK has trended up since 2008 to over 1.0% pa, as illustrated in Figure 1 below; and
- Ofgem has maintained a significant wedge in its OE targets above UK productivity growth. For example, Ofgem's headline[3] OE targets for gas networks have been relatively flat over time (marginally increasing from 0.85%[4] at RIIO-GD1 / T1 to 1.0%[5] at RIIO-GD2 / T2), rather than reducing to reflect the pattern seen across the economy. Further, in absolute terms these targets are well above the near-zero productivity growth observable for the UK economy.

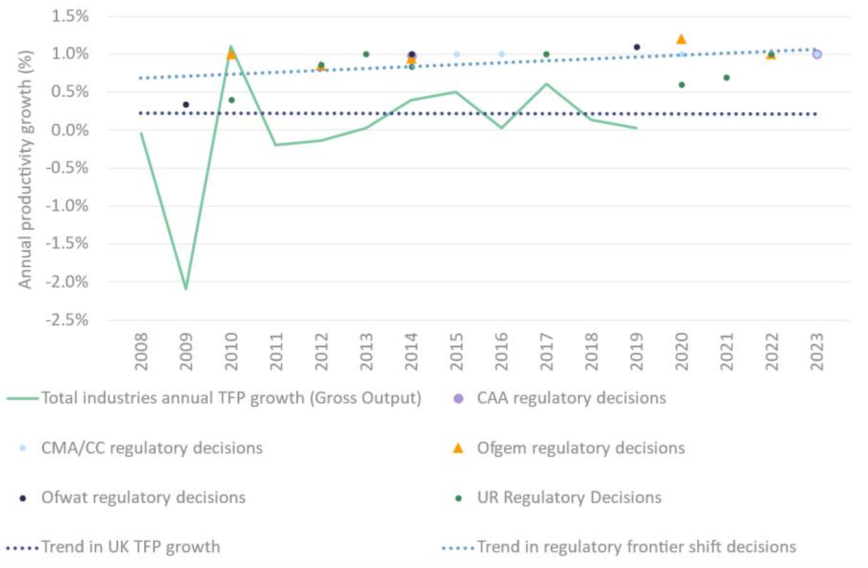
---

3 *For illustrative purposes we present the 'headline' OE target by taking a simple average of the opex target and capex / repex target. This is consistent with Ofgem's terminology used at RIIO-2, for example see: 'RIIO-2 Final Determinations - Core Document (REVISED).' Ofgem (February 2021); para. 5.27.*

4 *The opex challenge was set at 1.0% and the capex / repex challenge was set at 0.70%. See: 'RIIO-T1/GD1: Final Proposals – Real price effects and ongoing efficiency appendix.' Ofgem (December 2012); para. 3.3.*

5 *The opex challenge was set at 1.05% and the capex / repex challenge was set at 0.95%. This was following the removal of the 0.2% innovation uplift at the CMA appeals. See: 'Cadent Gas Limited, National Grid Electricity Transmission plc, National Grid Gas plc, Northern Gas Networks Limited, Scottish Hydro Electric Transmission plc, Southern Gas Networks plc and Scotland Gas Networks plc, SP Transmission plc, Wales & West Utilities Limited vs the Gas and Electricity Markets Authority: Final determination Volume 2B: Joined Grounds B, C and D.' CMA (October 2021); para. 7.4 and para. 7.867.*

Figure 1: Stagnant UK productivity growth and increasing regulatory OE decisions, post-2008



Source: Economic Insight analysis of EU KLEMS data and regulator decisions.

Intuitively, if one applied a consistent method for setting OE over time, then (irrespective of whether one thought regulated companies might out- or underperform UK productivity growth overall), the trend in those decisions should still *broadly reflect* the trend in UK productivity performance. The fact that it does not thus reflects changes in methods by regulators over time.

Secondly, we observe that regulator decisions have not adequately considered, or accounted for, the problem of volatility in parts of the underlying productivity data. It is vital that regulator-set OE targets reflect the *'most likely'* productivity gains companies can achieve. As such, they should not be driven by volatility in underlying productivity data. A failure to take this into account may result in OE targets being due to material changes in inputs our outputs within comparator industries *that are not indicative of underlying productivity performance.* We have identified two areas where this volatility is particularly problematic:

- **Using VA, rather than GO, as the productivity metric.** The omission of intermediate outputs under VA, for ease of calculation, increases the volatility of the metric. We observe this in our own analysis, but it is also well established in the academic literature.

- **Using an overly narrow set of comparator industries**. This is because industry level estimates of productivity are inherently more volatile than more aggregated measures. For example, this might intuitively arise due to spikes in capex in heavy investment periods for an individual industry, which get *'averaged out'* if one measures productivity across a greater number of industries, or for the economy overall. Similarly, smaller sample sizes may inherently reduce estimation accuracy of the underlying productivity data for some industries.

CEPA's methodology at RIIO-2 suffers from both of the above issues. Indeed, we find that, when we perform a 'complete update'[6] of CEPA's methodology from RIIO-2 (using the most recent EU KLEMS data and time periods), starkly different OE estimates are implied for RIIO-3.

In particular, the 'complete update' of CEPA's results indicates an OE target at RIIO-3 of between -1.1% and 0.5%; a dramatic change from CEPA's RIIO-2 range of 0.2% to 1.0%. We conclude that this arises from both CEPA's use of a narrow comparator set and the VA metric. The instability of CEPA's results indicates a critical need to give more explicit consideration to the underlying data (and method choices) being used going forwards.

1B. Now is the right time to apply a fresh perspective to OE and establish a long-standing approach, which is guided by clear principles, and applied consistently going forward

In light of the above, it is now essential to step back and think afresh about OE. Most importantly, a principles-driven approach should be transparently established. The objective being that this can then be consistently applied over time at future price controls. The benefit of this is that future variations in regulator-set OE would then be more likely to reflect changes in underlying productivity performance for regulated industries, rather than being due to detailed changes in method choices or data, which may be subjective; opaque; and not well-evidenced.

To the above end, we recommend the approach for estimating OE at RIIO-3 be guided by the following principles: (i) the benchmarking approach should be transparent and robust; (ii) the relevance of UK productivity performance to gas networks should be considered; (iii) post-estimation adjustments to the range derived from benchmarking should be avoided; and (iv) point estimates from any benchmarked range should *generally* be taken from values *'towards the middle'* of that range.

The benchmarking approach should be transparent and robust

Benchmarking requires choices to be made as to the: (i) **productivity metric** to be used (VA or GO; and TFP, or TFP combined with labour); (ii) **comparator industries;** and (iii) **time periods.** These choices should be: based on robust theory and evidence; transparent; and consistent across regulatory decisions.

Productivity metric choices

With regard to the choice of productivity metric, the analysis and evidence developed in this report supports the use of TFP for all costs (rather than some combination of TFP and partial factor metrics, such as labour productivity for opex). This is because a *'mixing and matching'* of TFP and partial factor metrics ignores the fact that the

[6] *By this we mean we use exactly the same comparator industries as CEPA but update: (i) the dataset to the latest EU KLEMS release; and (ii) the time period to include the full set of years now available (i.e. 1995-2019). We provide more detail in Annex 4.*

comparator industries used in any benchmarking are themselves free to substitute between labour and capital. They should, therefore, logically be using the optimal mix of these inputs to maximise their TFP. Hence, by setting gas networks an overall OE challenge that includes some weighting of *'benchmarked TFP'* and *'labour productivity'*, there is an incorrect implicit assumption that productivity can be boosted through the use of more labour. Or, put another way, it assumes the comparators being used to generate the TFP benchmark were not using an optimal mix of inputs. Notwithstanding this, we find that labour accounts for less than half of gas networks' opex[7].

As to the choice between VA and GO, a range of theory and empirical evidence supports the use of GO. Specifically; (i) the academic literature establishes (and the OECD recommends) GO as a conceptually superior measure; (ii) GO accounts for intermediate inputs, which are a material proportion of costs for gas networks (c. 50% of controllable opex); (iii) at an individual industry level, GO is a better measure of productivity than VA (being less volatile, but also avoiding a potential upwards bias that arises for VA).

Comparator industry choices

Clear and transparent criteria should be used to identify comparator industries. We have used the following criteria: (i) similarity of activities; (ii) extent of competition; (iii) similarity of scope to benefit from economies of scale. We consider these to be non-contentious. However, it is important not to place undue weight on the apparent similarity of activities, particularly if the assessment of that is relatively superficial. This is because it may result in *'too few'* comparators being included, which may make OE unstable under future updates, due to the underlying data volatility issue, and in turn reduce confidence that the changes reflect shifts in underlying productivity potential.

Our approach weights our criteria equally and is also data driven. As such, it avoids the *'too few comparators'* issue, whilst also ensuring we avoid including comparators that are not sufficiently representative of OE for gas networks.

Time periods

The time periods over which productivity growth should be benchmarked to estimate OE should largely be driven by internal consistency. That is to say, across RIIO-3, the regulatory method should be clear on: (i) the *'overall question'* being answered (e.g. *'what should allowed revenues be so as to reflect the efficient costs and risks faced by companies and customers over the RIIO-3 period?'*); (ii) when drawing on historical data to set regulatory parameters, what period is most appropriate to answering the question; and (iii) how consistency is achieved across the regulatory building blocks.

The answer to (i) should ideally not change across price controls (i.e. one is either seeking to set optimal determinations for the regulatory period, or something longer term, but in either case, the view taken at one price control should not be revisited at future ones). In turn, that should mean there is no sound basis to continually revisit the

7 *This is excluding contractor labour, which is an intermediate input, not a labour cost.*

time horizon debate for OE.[8] We also consider that it is preferable that the choice of time period allows for: (i) utilisation of as much data as possible, to reduce the chance that results arise out of a 'fluke' in the chosen years of data; and (ii) use of full business cycles, as productivity is generally higher in upswings in the business cycle; and (iii) the structural break in UK productivity growth, which is addressed below.

The relevance of UK productivity growth to gas networks must be considered

To differing degrees, sectoral regulators have argued to attach less weight to the post-2008 time period (the period of near-zero UK productivity growth). Thus, contrary to the above principle, the time horizon estimation choice made by regulators has varied across determinations. Regulators have argued that this is appropriate, as the factors causing the slowdown may not apply, or apply less strongly, to regulated industries.[9] In our report, we examine this topic in detail. We have: (i) developed extensive evidence on the specific factors driving the productivity growth slowdown (including a comprehensive review of academic literature, and drawing on a survey of the UK's leading academic experts on productivity);[10] and (ii) evaluated the extent to which those factors apply to gas networks. We find limited reasons to suppose gas networks are materially shielded from the causal factors of the slowdown:

- **The main factors causing the UK productivity growth slowdown are largely economy-wide and are unlikely to fully unwind over RIIO-3.** Evidence shows the key causal factors of the slowdown are insufficiency of: (i) investment; (ii) infrastructure quality; (iii) human capital quality; and (iv) management quality.
- **Regulation is unlikely to mitigate the impact on gas networks of the factors causing the slowdown.** Regulation can only credibly mitigate the problem of underinvestment. However, data shows that, even on this issue, regulated industries are not insulated. As there are no strong reasons to believe regulation mitigates any of the other factors, its overall mitigating impact must be negligible.

In light of this evidence, we conclude that the plausible lower bound for OE is given by productivity growth in the most recent business cycle (2010-2019). This is because we think that productivity growth is unlikely to deteriorate any further, so a continuation of the recent present provides a plausible lower bound. The plausible upper bound is

---

8 To clarify, clearly specific time periods used in benchmarking would change, as new data became available at each price control. However, the underlying matter of 'the relevant estimation window' should not change once the answers to (i) and (ii) are reached. The decision by regulators to attach lower weight to the recent past, following the financial crisis, is one example of this (where, as explained in footnote 8) that issue would be better addressed through comparator selection.

9 In our view, this conflates: (i) the key in principle question of what the appropriate estimation time horizon is; with (ii) the identification of appropriate comparators. Namely, the implication of regulator arguments on this issue is that the causal factors driving the productivity growth of the comparators are less relevant to regulated industries (in this case, gas networks). Moreover, to date, the arguments forwarded to propose that the pervasive slowdown in productivity growth across the UK (and other Western economies) have less impact on regulated industries are largely based on hypotheses and are not evidence-based.

10 This survey was part of a wider academic research exercise, the results of which are contained in: 'The UK productivity puzzle: A survey of the literature and expert views.' Williams, S.; Glass, A.; Matos, M.; Elder, T.; and Arnett, D. (January 2024). This wider research exercise was not undertaken on behalf of (nor funded by) any clients. Participants in our research did so of their own volition and without any financial incentive. The research was not undertaken for the purpose of making regulatory submissions and was an academic endeavor on the part of the listed authors.

provided by a weighted average of: (i) 1995-2019; and (ii) 1970-2007. Our view is that this provides a likely upper bound for OE at RIIO-3 because we think it unlikely that the structural break in productivity growth will *fully unwind* over RIIO-3. By averaging between (i) and (ii) above, however, we allow for a *partial unwinding*.

Post-benchmarking adjustments should be avoided

There are various reasons why one might consider making post-estimation adjustments to a benchmarked OE estimate (e.g. that TFP includes catch-up- and economies of scale-related gains; the question of to what degree embodied change is captured; etc). Our report considers such issues in detail. However, we find that *'at best'* the directional impact of each issue can be identified and there is no reliable way to quantify their impact on OE. Therefore, one cannot determine whether the appropriate net impact of these factors would be that the benchmarked OE is under-/overstated. Thus, one cannot reliably determine whether a net upwards or downwards adjustment is appropriate. We thus do not recommend post-benchmarking adjustments to OE.

We are further concerned that the magnitude of such adjustments can be (and has been) so material as to call into question the validity of the benchmarking method in the first place. For example, at RIIO-2 Ofgem made two post-benchmarking uplifts to the OE estimate for transmission and gas distribution networks. This included a 0.2 percentage point adjustment for the innovation fund; and adjustments to add further stretch (for example, placing less weight on the wider productivity growth slowdown of recent years). Together, this accounted for approximately 50% of Ofgem's final OE estimate.

Point estimates from any benchmarked range should generally be taken from values 'towards the middle' of that range

As a point of principle, we consider best practice should be to derive any OE point estimate from towards the middle of any range derived from benchmarking. This reflects the inherent uncertainty as to the *'true'* value of OE, where it cannot be observed. It would be appropriate to depart from this if there were compelling evidence to the contrary (on a case-by-case basis), but we do not observe such evidence in the present case. Again, this principle should help drive consistency over time and avoid accusations of cherry picking in either direction.

1C. Our recommended range for OE at RIIO-3

Based on our method summarised above (and described in more detail in the remainder of this report), we consider it highly likely that OE for gas networks at RIIO-3 will lie in the range of 0.2% and 0.8% (with a midpoint of 0.5%). This is our recommended range. We do not advocate any particular point estimate within this range, but note the principle discussed above that it is typically appropriate to draw on values towards the middle of the range. Of specific relevance to OE at RIIO-3, we note:

- Academic survey evidence shows most experts expect UK productivity growth to be 0.5% pa or below over the next five years.
- Academic survey evidence further shows that expert academics do not expect the energy industry to outperform the UK, with regards to productivity growth.
- When using CEPA's benchmarking approach at RIIO-2, with the benefit of more up-to-date EU KLEMS data, the upper bound of the resultant range for OE is 0.5%.

Finally, whatever the determined number for OE, it is critical to note that (as explained by Professor Anthony Glass – see Annex 10) this will be *inclusive* of productivity gains realised through improvements in quality / output. Therefore, suppose one's best view of OE for RIIO-3 was 0.5% pa. This would mean that, if the entirety of this was applied to gas networks' costs, but those companies were additionally tasked with making improvements in quality / output, there would be a double-count. This does not imply the need for any post-benchmarking adjustment to OE. Rather, it implies the estimate should be *allocated* between reduced costs / improved quality / output.

The remainder of this report covers: (i) our approach to comparator selection in Chapter 2; (ii) the wider evidence around UK productivity growth in Chapter 3; (iii) adjustments to the benchmarking range in Chapter 4; and (iv) technical annexes.

2 Comparator analysis

Estimates of OE for regulated industries are typically derived from productivity growth data, using a benchmarking approach. At a high level, this involves selecting a set of sectors that can be considered comparable to the regulated industry and calculating the average annual productivity growth across them over a certain period of time. This average is then used as the benchmark to set the OE challenge.

We apply this same approach in this report, drawing primarily from EU KLEMS data. We detail our method by outlining our preferred choices for the following key analytical decisions: (i) the productivity measure; (ii) the time period; and (iii) the comparator sectors.[11] We then present the resultant OE estimates for our '*recommended range*'.

2A. Choice of productivity measure

In this chapter, we outline our choice of productivity measure and discuss how OE differs from measured productivity growth. We conclude that the following analytical choices are most suitable for gas networks:

- OE estimates should be based on **TFP**, rather than partial factor (e.g. labour) productivity. This is because TFP uses all measurable factors of production that it is possible to include. It is therefore more reflective of the costs of gas networks.
- OE estimates should be based on the **GO measure of productivity,** because it more accurately measures changes in productivity over time and across industries than VA; the alternative measure of productivity growth.

OE estimates should be based on TFP, rather than labour productivity

Productivity gains made from OE are most frequently assessed through TFP analysis. TFP is a measure that "*captures changes in performance attributable to increased physical production of outputs relative to inputs*".[12] TFP captures all the measurable factors of production that it is possible to include. Inputs are typically measured as capital and labour but can also include intermediate inputs if TFP is measured on a GO, rather than VA, basis (we explain this in more detail in the next subsection). Output is usually a measure of aggregate economic output. TFP thus represents the change in output that cannot be explained by changes in the quantity of capital and labour (and, in the case of GO, intermediate inputs).[13]

Our view is that TFP is the most appropriate measure upon which to base the OE challenge for all costs. This is because TFP, as stated above, reflects all the inputs that are relevant to the costs of gas networks (and not just labour). Consistent with this, we

---

11 *We use 'sectors' and 'industries' interchangeably throughout this report.*
12 *'Regulatory Price Performance, Excess Cost Indexes and Profitability: How Effective is Price Cap Regulation in the Water Industry?' Maziotis, A; Saal, D; Thanassoulis, E (September 2009); page 5.*
13 *And so, TFP growth is considered to comprise of intangible factors, such as technological change; R&D; and synergies.*

note that Ofgem solely relied on TFP to set the OE challenge for: (i) capex & repex at RIIO-T2 & GD2; and (ii) for totex at RIIO-ED2.[14]

'Opex for gas networks contains a significant proportion of non-labour costs.'

It has previously been argued by Ofgem that some weight should be placed on labour productivity measures when setting an opex specific OE challenge. For example, Ofgem placed some weight on labour productivity when setting the opex challenge at RIIO-T2 / GD2.[15] However, it is somewhat unclear *'how much'* weight was placed on this. Contrary to this, our view is that the OE challenge for all costs (including opex) should be based *solely* on TFP. This is for the following reasons:

- The *'mixing and matching'* of TFP and (partial factor) labour productivity ignores the fact that the comparator industries used in the benchmarking are free to substitute between labour and capital. Thus, they should logically be using the optimal mix to maximise their TFP. Therefore, basing the OE challenge for gas networks on some weighted average of TFP, and TFP and labour (for opex), erroneously assumes that further gains can be made by using more labour. Put another way, it assumes the comparators were not using an optimal mix of labour and capital. This is not conceptually sound, as economic theory dictates firms will profit maximise. Therefore, an OE challenge based on a TFP benchmark already reflects *the maximum productivity achievable using the optimal mix of inputs.*

- Opex for gas networks predominantly consists of non-labour costs. For example, our analysis of data provided to us by the ENA indicates that staff costs make up a minority (46%) of their opex, on average[16]. This is consistent with NGN's statement at the RIIO-2 energy appeals, where it reported that staff costs only made up 49% of its opex.[17]

The EU KLEMS dataset (which provides a measure of TFP) has typically been used to arrive at estimates for the OE challenge, including at RIIO-2; PR19; and the CMA PR19 redeterminations. In this report, our OE estimates are also based on EU KLEMS TFP data, using the latest release available.[18]

Gross output is the appropriate measure of TFP, rather than value added, for the purpose of setting OE

The GO and VA measures are *distinguished by their approach to intermediate inputs:* the former includes them, while the latter excludes them. Intermediate inputs are those factors of production that are produced and transformed during (or used up by) the production process.[19] This is by contrast to *primary* inputs, comprised of labour and

14 *At RIIO-ED1, Ofgem did not conduct its own analysis. Instead, it accepted company submissions.*
15 *'Energy appeals: Final determination, volume 2B: Joined Grounds B, C and D.' CMA (October 2021); table 7-2.*
16 *This is excluding contractor labour, which is an intermediate input, not a labour cost.*
17 *'Energy appeals: Final determination, volume 2B: Joined Grounds B, C and D.' CMA (October 2021); para. 7.170.*
18 *This comprises TFP data from both the NACE II (1995-2019) and NACE I databases (1970-2007). We use the latest version of the EU KLEMS NACE II database, which was released in February 2023.*
19 *For example, the intermediate inputs of a car manufacturer may include: (i) materials such as steel used in the production of the car; (ii) energy such as the electricity used to run the machinery that builds the vehicles; and (iii) services such as leasing a truck to transport the final vehicles produced from the factory to the manufacturer's retail outlet.*

capital, which are not used up by the production process.[20] The exclusion of intermediate inputs from VA produces different productivity values and when selecting which of GO or VA to use, one must assess the advantages and disadvantages of each in general, *but also relative to the objectives at hand.*

We first outline: (i) how TFP is most often calculated in practice; (ii) which measure is favoured by regulators; and (iii) which measure is favoured by the CMA. We then explore the arguments for using GO or VA in detail.

In practice, TFP data is more frequently presented using the VA methodology. The often-cited reason is that it is easier to calculate than GO. Specifically, both the EU KLEMS and ONS datasets use a VA methodology. However, the EU KLEMS dataset also reports GO figures (which, in the case of the UK, are based on the VA measures produced by the ONS, which are then adjusted, so as to be on a GO basis).

The regulatory precedent on GO vs VA has been inconclusive; different regulators have followed different approaches and the CMA has not taken a view on which is '*best'.* At RIIO-2, Ofgem (following the advice of its consultants, CEPA) used both VA and GO measures to inform its OE challenge. CEPA argued this was because "*no consistent expert view has emerged on which one should be preferred.*"[21] At the energy appeals, the main practical or conceptual[22] reason given by Ofgem in favour of VA was that it considers there to be *"practical difficulties in estimating GO".*[23]

The CMA's historic approach appears consistent with choosing the metric that is most appropriate for the specific sector being regulated. The CMA at the PR19 redeterminations (whereby the authority was making its own judgement on each element of the price control) concluded that *"we therefore decide to focus on the gross output measure but give some qualitative weight to the value added metric."*[24] Whereas, under the energy appeals (whereby the authority was tasked with considering whether GEMA was *'wrong'*), the CMA concluded that: *"the appropriate weighting to attach to the VA productivity measure and the GO measure is a matter of regulatory judgement and different regulators can take different views on this topic."*[25] The CMA also stated that "*for example, if a regulator has concerns about whether the assumptions underlying the calculation of VA measures are applicable to the industry it is regulating it would be appropriate to place less weight on VA measures*."[26]

---

20 'Measuring Productivity: Measurement of Aggregate and Industry-Level Productivity Growth.' OECD (2021).

21 'RIIO-GD2 and T2: Cost Assessment – Frontier shift methodology paper.' CEPA (May 2020)

22 The discussion was complicated by the fact that it was unclear how much weight Ofgem had placed on VA in the first place. Furthermore, many of the other arguments related to either regulatory precedent or whether the VA figures were achievable for gas networks. As our method is based on best practice with regards to a ground up approach to benchmarking, we do not consider it necessary to address these arguments.

23 'Energy appeals: Final determination, volume 2B: Joined Grounds B, C and D.' CMA (October 2021); para. 7.139.

24 'Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations: Final Report.' CMA (March 2021); para. 4.545.

25 'Energy appeals: Final determination, volume 2B: Joined Grounds B, C and D.' CMA (October 2021); para. 7.146.

26 'Energy appeals: Final determination, volume 2B: Joined Grounds B, C and D.' CMA (October 2021); para. 7.148.

In the remainder of this section we assess the advantages and disadvantages of using GO or VA to inform the OE challenge for gas networks. We conclude that there is a **clear rationale for using GO** at RIIO-3 because:

- Although removing intermediate goods in the VA measure may offer a more precise measure of productivity, it offers one that is less accurate. We consider that using VA to avoid *"practical difficulties in estimating GO"* is unintuitive, as VA is an intrinsically flawed measure of the productivity gains achievable by gas networks in the first place (we discuss why this is the case below).

- Removing intermediate goods in the VA measure leads to bias in the estimation of productivity. This is particularly relevant for gas networks, because their controllable opex contains a material amount of intermediate inputs (c. 50%).

- It is established in the literature (and recommended by the OECD) that the GO measure is preferred when looking at industry specific productivity, which is the context of benchmarking studies.

We elaborate on each of these reasons below.

Removing intermediate goods in the VA measure may offer a more precise measure of productivity, but one that is less accurate

The differences between VA and GO occur due to the exclusion of intermediate inputs from the VA measure. The commonly cited rationale for the exclusion of intermediate inputs is that they may vary greatly by industry and can be difficult to quantify[27], which can lead to measurement error if they are included (as in the GO measure). Removing them, as is the approach in the VA calculation, serves to shortcut around this issue: i.e. if intermediate inputs are removed, they cannot cause measurement error. This gives the theoretical benefit of *precision* to estimates of VA.

However, this is only a half-solution; the VA measure loses *accuracy* with regard to what we are ultimately interested in for regulatory purposes. The relevant measure of productivity for the estimation of OE is the efficiency with which firms turn <u>all</u> their inputs into outputs, which is captured by GO. VA falls short of this and only measures the productivity with which firms <u>add value</u> to their inputs. This can lead to OE being set at the incorrect level.

We illustrate these differences between VA and GO in the figure below, where the centre of the target represents the '*correct*' estimate of productivity i.e. the one that is most appropriate for the OE challenge. In our explanation, we refer to two key terms:

- **Accuracy**. An *accurate* measure of productivity in this context, is one that seeks to measure productivity appropriately for an OE challenge i.e. the measure is aiming for the centre of target (but may not consistently hit it, due to errors in the *precision* of estimation).

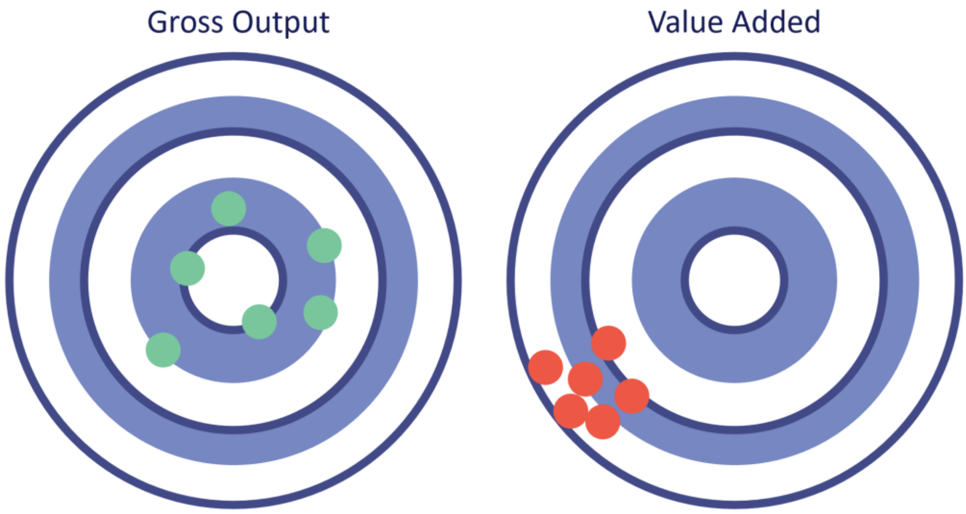
---

27 *'The quadratic approximation lemma and decompositions of superlative indexes.' Diewert, W. E. (December 2002).*

- **Precision**. A *precise* measure of productivity, is one where the productivity *estimates* are closely clustered around each other, since there is minimal error in the practical estimation of the metric, i.e. the measure will consistently hit the same point on the target, but this may not be the centre of the target.

Using these terms, GO prioritises *accuracy* over *precision*; while VA prioritises *precision* over *accuracy*. Hence, while productivity *estimated* by VA will closely approximate theoretical *true* VA productivity; VA productivity growth is intrinsically flawed, because it does not seek to measure what it needs to in order to inform an appropriate OE challenge for gas networks. The opposite is true for GO.

Figure 2: GO is a more accurate measure of productivity as should be measured when setting OE



Source: Economic Insight
Note: The middle of the target represents the 'correct' estimate of productivity. Estimates that are highly precise and accurate would be closely clustered together around the centre of the target. Accurate estimates are close to the 'correct' estimate at the centre, whereas precise estimates reliably get the same results, regardless of accuracy.

We consider that the problems of inaccuracy (with VA) are significant and we explore these in further detail below. By contrast, we consider that, with a sufficient number of data points (be it through a sufficient number of comparators, or utilisation of a longer time series of data), the issue of precision under GO can be mitigated.

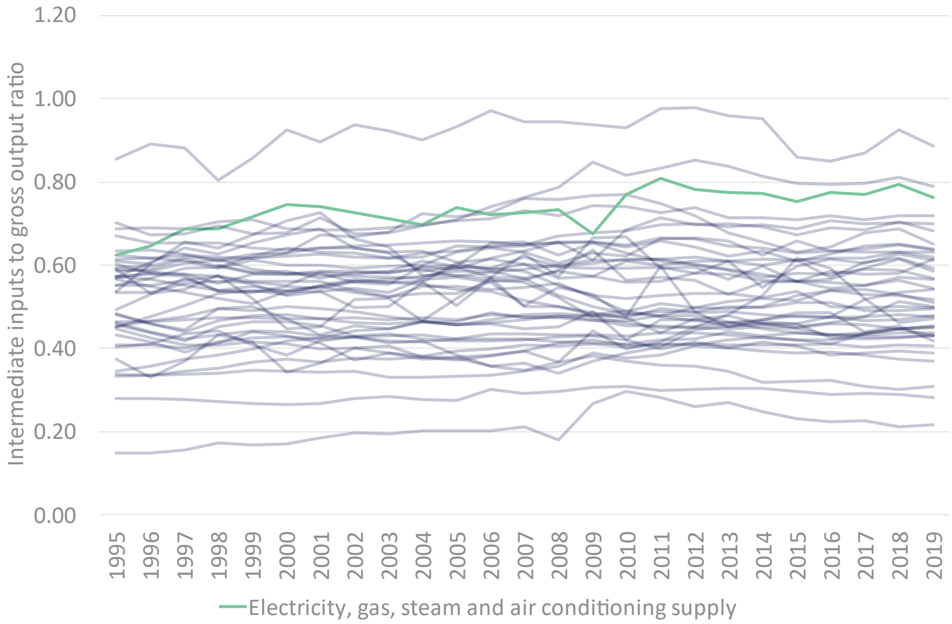
Removing intermediate goods in the VA measure leads to bias in the estimation of productivity

Improvements in productivity growth can arise from increases in efficiency in the use of intermediate inputs; and we therefore consider that it is *critical* to account for this when setting OE. As stated above, the VA approach removes all intermediate inputs from the estimation of productivity growth and is therefore prone to exclude an important source of economic growth (or decline). For this ommission, Diewert (2001)[28] criticises VA productivity growth as being systematically biased upwards.

[28] *'Productivity trends and determinants in Canada.' Diewert, E.W.; Department of Economics, University of British Columbia (2001).*

It is further established that this systematic upward bias in VA growth calculations disappears if the ratio of intermediate inputs to GO is constant across time and industries.[29] We test for this by examining the ratio of intermediate inputs to GO, as seen in the figure below. We find that this ratio varies significantly across time and industries. For '*Electricity, gas, steam and air conditioning supply'* specifically, this ratio grows from around 0.6 to 0.8, indicating an increase in the relative use of intermediate inputs over time. Put simply, in practice, the data is consistent with VA being systematically upwards biased (including for the purpose of setting OE for gas networks).

Figure 3: Intermediate inputs to gross output ratio across industries (1995-2019)



Source: Economic Insight analysis of EU KLEMS data.
Note: Lines in blue represent all other industries present in the EU KLEMS data.

In summary, while both measures of productivity growth are perhaps suitable for establishing whether productivity growth is increasing or decreasing, the VA measure is likely to overstate the extent of the change in the growth rate in comparison with the GO measure, *particularly for industries with a material utilisation of intermediate inputs*, such as gas distribution.[30]

Economic theory indicates that further bias in the VA measure will arise in a number of scenarios:

- If an industry experiences a change in its productivity of intermediate inputs, this will not be measured correctly in the VA approach. The GO approach acknowledges that intermediate inputs are a source of industry growth, offering a more complete picture of productivity improvements.

[29] *'Accounting for the Growth of Output.' Star, S.; American Economic Review (1974).*
[30] *'Measuring Productivity: Measurement of Aggregate and Industry-Level Productivity Growth.' OECD (2021).*

- If an industry experiences a change in the prices of intermediate inputs, leading to a change in the industry's output. This holds particular significance for gas networks in the current macroeconomic environment, which is characterised by significant fluctuations in intermediate input prices (such as energy). In the analogous circumstances of the 1970s energy crisis, Norsworthy and Jang (1992)[31], demonstrate how the significant changes in intermediate input prices reveal the above shortcoming of using VA, rather than GO, for productivity estimation. It is possible for the VA growth rate to grow, while the GO growth rate is constant if, for example, *the price of an intermediate input is falling.* This is because the cheaper intermediate input allows the industry to increase its output (all else constant). Under the VA approach (which does not measure intermediate inputs), industry productivity appears to have increased, because the industry is generating more output with the same labour and capital inputs.

- If an industry experiences a change in its input mix that leads to a change in the use of intermediate inputs. A typical example of this problem is when companies outsource their labour (a primary input) to more efficient contractors (an intermediate input). This may truly yield an increase in productivity (which will be picked up by the GO measure). However, this will be overstated by the VA measure, which registers the associated rise in output, following a reduction in labour inputs (but excludes the offsetting use of outsourced labour).

It is important to note that the comparator-based approach relies on valid comparisons of productivity growth, both over time and across industries. It is therefore critical that intermediate inputs are included in our measure of productivity if:

- Intermediate inputs are important in our industry of interest.

- The use of intermediate inputs varies across time and industries.

We note that Figure 3 demonstrates both of these characteristics. Furthermore, data provided by the ENA for the purpose of this study indicates that a material proportion of their controllable opex[32] is comprised of intermediate inputs, where the average across the companies is approximately 50%.[33]

It is established in the literature that the GO measure is preferred when looking at industry level productivity

[Value added is] 'not a good measure of technology shifts at the industry or firm level' – OECD Manual for Measuring Industry Productivity Growth.

While it is the case that the GO / VA measures of productivity yield very similar results at a national level, this is not the case when looking at the industry level. Thus, for the task at hand (where we make an industry specific assessment) the choice is important.

---

31 *'Empirical measurement and analysis of productivity and technological change: applications in high technology and service industries.' Norsworthy, J. R. and Jang, S. L.; Contributions to Economic Analysis Series (1992).*

32 *We use controllable opex primarily because it is the only cost category for which a detailed enough breakdown of costs is available to classify intermediate inputs.*

33 *We calculate this figure by allocating different costs as either intermediate input costs or primary input costs in Table 2.01 of company RIIO-2 Business Plan Data Tables for opex (in the case of National Gas the data was prepared to a data disclosure request as National Gas has different reporting requirements). Cost allocations were made following discussions and clarification questions with the companies, and economic theory. As cost categories are not fully atomised, we consider that this proportion could reasonably range between 48%-52%.*

Relatedly, Cobbold (2003) states that *"there are theoretical grounds for preferring the gross output approach, particularly at the industry level"*[34] and the OECD's Manual for Measuring Industry Productivity Growth (2021)[35] states that VA is *"not a good measure of technology shifts at the industry or firm level"*.

In the EU KLEMS data, we observe that **the VA measure of productivity growth is highly volatile at the industry level**. This can be seen for the '*Electricity, gas, steam and air conditioning supply*' industry, in the figure below, which at times changes by more than 40 percentage points between years. This volitility gives us less confidence in the use of VA for the present industry level study. We note that this is not inconsistent with the idea that *estimated* VA productivity is precise but not accurate. Instead, *true* VA productivity exhibits high volatility, and *estimated* VA productivity consistently captures this volatility.

VOLITILITY IN THE VALUE ADDED MEASURE AT AN INDUSTRY LEVEL IS PROBLEMATIC.

Figure 4: The VA measure of productivity growth is highly volatile at the industry level

Source: Econonic Insight analysis of EU KLEMS data

This is by clear contrast to the relative volatilities of the two measures at the national level, where they are more similar (although, VA does still exhibit greater extremes than GO), as seen in the figure below. From the data, we infer that the VA measure experiences additional bias from the exclusion of intermediate inputs at an industry level, rendering it unreliable for the purpose of setting OE under regulatory detrminations.

[34] *'A Comparison of Gross Output and Value-Added Methods of Productivity Estimation.' Cobbold. T (2003).*

[35] *'Measuring Productivity: Measurement of Aggregate and Industry-Level Productivity Growth.' OECD (2021).*

Figure 5: The VA measure of productivity growth performs better at the economy level

Source: Economic Insight analysis of EU KLEMS data.

2B. Choice of time periods

Estimates of OE are sensitive to the time period over which they are assessed, and therefore the choice of time period needs careful consideration.

Our view is that time period selection should be largely driven by **internal consistency.** This consideration is particularly significant for setting OE because productivity, growth, and equity returns are all correlated. It is therefore crucial that the time horizon used (and wider economic context assumed) for the purpose of setting other key elements of RIIO-3 is consistent with that assumed when setting OE.

There are several further considerations that should be taken into account when determining which time period(s) to use for setting the OE challenge. These are as follows:

- **Utilisation of the full data available.** Maximising the number of observations used in estimating OE reduces the risk of outliers affecting the results. EU KLEMS is comprised of two datasets: NACE I (1970-2007); and NACE II (1995-2019).

- **Full business cycles.** Because productivity is pro-cyclical, one is more likely to obtain a balanced OE estimate by ensuring any analysis includes a full *'peak-and-trough'* business cycle. Credible independent sources should be used to inform appropriate start and end dates for such cycles, to avoid the critique that chosen time periods are arbitrary. Annex 7 sets out our views of when the business cycles occur.

- **The structural break arising after the financial crisis.** There has been a structural break in UK productivity growth, which has flatlined since the financial crisis in 2008 (and shows no sign of returning to pre-crisis levels in the near-term). We explore the evidence for this in detail in Chapter 3. Since RIIO-2, EU KLEMS data has been released covering more recent years (up to 2019). This means it is possible to use a full business cycle's worth of data post-crisis (which our analysis in Annex 7 indicates is 2010 to 2020). Placing increased weight on more recent time periods may be more appropriate when determining OE on a forward-looking basis, given the clear persistence of depressed productivity growth performance in the UK (i.e. this is not an energy industry specific issue). Nonetheless, we also consider that it is helpful to include data from periods that allow us to show what OE *could* be if the structural break in productivity growth unwinds (to some degree) over the course of RIIO-3.

Trade-offs between the above considerations must be made when determining which period to analyse. For example, if new data were to be released two years after the end of a given business cycle, while we would want to use this new data so that our analysis benefits from an increase in the number of observations, we would also have to factor in how this would add an 'incomplete' business cycle to the data (and the associated drawbacks of doing this).

With these considerations in mind, we consider two different date ranges (each containing multiple time periods) for our analysis, which are as follows:

Recommended range

This is the range, which we think it is likely for OE to be in at RIIO-3. This range contains two time periods:

- **2010-2019 (EU KLEMS NACE II).** This provides what we consider to be the plausible lower bound of productivity growth. This is because we think it is unlikely that productivity growth will deteriorate further (i.e. the recent past provides a plausible lower bound), based on our evidence in Chapter 3. This time period includes nearly all of the most recent business cycle (which we find to be 2010-2020). We note that TFP data is unavailable for 2020 within the NACE II database. As a result, we consider that the OE estimates for this period may be biased upwards, given the final year of poor economic performance in the UK is not captured in this period.

- **Weighted average of: (i) 1995-2019 (EU KLEMS NACE II); and (ii) 1970-2007 (EU KLEMS NACE I)**. In addition to the most recent business cycle, we include an estimate that aims to provide a long-term view. We use this period for two key reasons. Firstly, it utilises the largest possible sample size by capturing the entire period for which data is available. Using the largest possible sample is particularly pertinent due to the limited number of total observations available. Secondly, this period implicitly allows for some (but not full) unwinding of the productivity growth structural break over RIIO-3. This is because it balances the low productivity growth seen post financial crisis against higher productivity performance in the more distant past. It does this by including periods from both before and after the crisis. We think this provides a likely upper bound for OE as the evidence from Chapter 3 indicates that anything more than *partial* unwinding of the productivity puzzle over RIIO-3 is unlikely. We are unable to combine the NACE I and NACE II databases, because the data is recorded differently in overlapping years. We have therefore combined estimates from the two databases by calculating a weighted average.[36]

Sensitivity analysis

We also conduct sensitivity analyses, which (in addition to the three time periods discussed above), also contains the 1992-2007 (EU KLEMS NACE I) time period. This contains the majority of the final business cycle before the financial crisis and serves as an alternative indicator to the 1970-2007 period for what productivity growth could be if the structural break were to fully unwind over RIIO-3. We do not recommend using the results of this sensitivity analysis (or any of our other sensitivities) to inform the OE challenge. This is because the sensitivities use different assumptions and/or time periods to those we recommend and are only included to test the robustness of our results. The results of these sensitivity checks can be found in Annex 3.

2C. Choice of comparators

A robust choice of comparators is critical to ensuring an accurate OE challenge. Therefore, in selecting the comparators for our analysis, we use a clear and data driven selection process that is consistent with best practice. In this section we outline: (i) the criteria used to determine our comparators (and why these matter); (ii) that TFP includes multiple types of efficiency beyond OE, which informs our chosen criteria; (iii) the application of these criteria to choose our comparators; (iv) our resulting preferred set of comparators.

[36] *We weight each period by the number of years it contains. To understand why this is necessary, consider the following example. Suppose we had two time periods: (a) and (b). Suppose time period (a) contains a single year (this is the year 2010), and time period (b) contains 10 years (these are the years 2000-2009). If we take a simple average of the average number from time period (a) and the average number from time period (b), then we are effectively placing a weight of 50% on the year 2010 and a much smaller weight on each of the individual years in the period 2000-2009. Assuming that the data from each time period is equally reliable, then there is no reason to place a higher weight on 2010 than the individual years in time period (b). It would be more logical to weight each year equally by calculating a weighted average. The same logic applies to two larger time periods.*

Criteria

Best practice states that comparators should conform to three key criteria, which are relatively well established:

- **Criterion 1: Similarity of activities being undertaken.** To ensure that the parallels drawn between the comparators and gas networks are reasonable, it is important that both undertake similar activities. When activities are similar between firms, one would expect productivity gains to be similar. This is because the comparators will likely use similar processes and technology to gas networks. Similar activities we have considered include: (i) operation and maintenance of a complex network; and (ii) the construction of major infrastructure.

- **Criterion 2: Competitiveness of industry**. Using comparators that operate in competitive industries means that TFP growth is more likely to have been primarily driven by OE; and will be less driven by catch-up efficiency (we explain the reasoning behind this in the next subsection). Thus, by focusing on industries that are 'more competitive', this should allow us to somewhat mitigate (but not remove entirely) the *overstatement of OE* that arises from 'catch-up' efficiency being included in any TFP figures.

- **Criterion 3: Extent of scale effects.** TFP includes productivity gains achieved through economies of scale (we explain the reasoning behind this in the next subsection). It is therefore important that comparators have a similar scope for scale-related gains to the gas networks. This is to ensure that TFP estimates more accurately reflect achievable OE. There are two broad ways in which the comparator choice can be used to mitigate against this:

  (i) **Criterion 3a: Fixed costs.** We would expect there to be a high correlation between the extent of fixed costs in an industry and the extent of scale effects. Hence, having a similar proportion of fixed costs to the gas networks is an important consideration when selecting comparators. Selecting industries with very different proportions of fixed costs to the gas networks could either over- or understate the scope for OE.

  (ii) **Criteria 3b and 3c: Capital growth and output growth.** Efficiencies arising from scale effects vary over time, in part because they vary with growth rates. For example, for a given level of fixed cost, a faster-growing firm benefits more from economies of scale than a slower-growing firm. Hence, comparators that exhibit similar growth rates over time to the gas networks further allows us to ensure that scale-related gains are likely to be similar over the relevant time period.

It is important not to place undue weight on the apparent similarity of activities (Criterion 1), particularly if the assessment of that is relatively superficial. This is because it may result in *'too few'* comparators being included as similarity is assessed in such a narrow way. Due to the underlying data volatility issue, this may make OE unstable under future updates, reducing confidence that they reflect changes in underlying productivity potential.

Using a broader range of factors (such as Criteria 2 and 3) to assess similarity allows the casting of a wider net, ensuring that more nuanced aspects of similarity in comparator industries are accounted for (that are not clear from a surface level assessment of similarity).

We discuss why Criteria 2 and 3 are vital to consider, over and above Criteria 1, in more detail in the following subsection.

TFP measures other types of efficiency beyond OE

It is established practice for TFP metrics to be used to derive estimates of OE. However, it is important to note that TFP may include or exclude aspects of productivity other than OE.[37]

Whilst OE represents the efficiency improvements that it is possible for even the most efficient firms to make over a period of time, (as above) TFP merely measures the change in the quantity of outputs, relative to a change in the quantity of inputs (i.e. the change in outputs that cannot be explained by a change in inputs). Due to TFP's broader definition, there is debate over what is captured within the metric and the most appropriate way to interpret and apply it for regulatory price control setting purposes. The key issues are as follows:

- TFP captures multiple efficiency savings (**catch-up efficiency** and **economies of scale**[38]).
- The extent to which TFP captures **embodied technical change** is unclear.

We consider the relevance of these issues below.

TFP captures multiple efficiency savings

TFP is a measure of *all* efficiency improvements that have been made. Although OE is one way that a firm could achieve TFP growth, it is also possible to achieve an improvement in TFP through other efficiency improvements. These include:

- **Catch-up efficiency**. TFP estimates also include catch-up gains, which are distinct from OE gains. If a firm, or firms, within an industry are not already operating at the efficiency frontier, TFP growth can be achieved via a firm 'catching-up' to the frontier. Catch-up efficiency will be present for *all industries* to some extent, as none are perfectly efficient (i.e. no market is perfectly competitive), meaning that there will always be some firms that are operating behind the frontier. The implication of this is that comparator choices should take the competitiveness of industries into account (whilst recognising the *intrinsic overstatement* of OE derived from TFP estimates that nonetheless must persist).

[37] Note that, as discussed in Section 2A, TFP is preferred to labour productivity. Additionally, labour productivity is subject to similar issues described in this subsection.
[38] This is what Criteria 2 and 3 are designed to address.

- **Economies of scale**. These occur in scenarios where unit costs rise or fall, depending on whether a firm's output volume is increasing or decreasing. If an industry benefits from economies of scale, then an increase in inputs would lead to a *more than* proportionate increase in outputs, as the unit costs of producing the output would fall. This would show an improvement in TFP growth. However, it would not be caused by an outward shift in the production frontier (i.e. it would not be equivalent to OE). Whether TFP estimates will over- or understate OE for this reason depends on whether the scope for gas networks to make gains from economies of scale is greater, or smaller, than for the comparators used. The implication of this is that similarity in ability to reap scale economy benefits should be a consideration in selecting comparators for gas networks.

We address this issue of multiple efficiency savings using Criteria 2 and 3 respectively. We also explore them in further detail in our sensitivity analysis[39] (in Annex 3) and our discussion of possible post-benchmarking adjustments to the OE range in Section 4A.

The extent to which TFP captures embodied technical change is unclear

When measuring OE, it is important that both embodied and disembodied technological change are included, in order that the full scope for productivity gains is captured:

- **Embodied** technological change relates to productivity gains generated from the use of <u>*new*</u> technology and assets.

- **Disembodied** technological change captures gains from the use of <u>*existing*</u> technology and assets.

Productivity metrics include disembodied change, but there is uncertainty and debate (including within the academic literature, as well as within financial institutions and government agencies) as to the extent to which they include embodied change. We discuss the extent to which TFP estimates include embodied change in detail in Annex 6 and conclude that: (i) the exact *proportion* of embodied change included in TFP is unclear, but the evidence suggests that it is included to *some degree*; (ii) the proportion of embodied change included in TFP varies by industry.

Accounting for embodied change in our benchmarking analysis is therefore more complicated than catch-up efficiency and economies of scale. With catch-up efficiency and economies of scale, we need only worry about the extent to which gas networks have similar levels of catch-up efficiency and economies of scale to the comparator industries. By contrast, with embodied change, we must also concern ourselves with the extent to which the TFP data of individual comparators contains their full scope to benefit from embodied change. Conceptually, therefore, this issue might not be entirely addressed through comparator selection.

On this basis, we do not include a criterion in our main comparator selection (that creates our '*recommended range*') and instead:

---

[39] *We do not support using the results of these sensitivities (or any of our other sensitivities) for informing the RIIO-3 OE challenge. This is because the sensitivities are based on assumptions that we do not recommend and are only used to test the robustness of our results.*

- include criteria to assess embodied change in our sensitivity analysis[40] (in Annex 3); and
- consider the evidence for post-benchmarking adjustments to our estimated OE range to account for embodied change (in Section 4B).

Assessment against criteria

We have sought to apply our evaluation criteria transparently, so as to arrive at an objective view as to the appropriate comparators. In the following passages, we set out how we have applied these in practice. Note, before assessing any comparators against our set of criteria, we first filtered all the industries down to a set that contained just: (i) those previously used by Ofgem and CEPA at RIIO-2; and (ii) any further industries that we consider could share similar characteristics to gas networks.

Criterion 1: Similarity of activities being undertaken

For Criterion 1, we undertook a qualitative assessment of the extent to which we considered the industry to share similar activities to those of gas networks. For those comparators we assessed, our ranking system is as follows:

- **Green**. These correspond to industries that we consider to be either identical (or very similar) to the gas networks. This includes: (i) the energy sector, '*Electricity, gas, steam and air conditioning supply*';[41] (ii) the water sector, '*Water supply; sewerage, waste management and remediation activities*'; and (iii) the combination of these two sectors, '*Electricity, gas, steam; water supply, sewerage, waste management*'.
- **Amber**. These correspond to the vast majority of industries assessed, as we consider most of these industries to share *some* activities with gas networks, but only to a degree. For instance, the construction sector is clearly involved in *'the construction of major infrastructure*' but does not '*operate and maintain a complex network'*.
- **Red**. These are sectors that we consider to be very different to the gas networks. For instance, we consider neither '*Financial and insurance activities*' nor '*Arts, entertainment and recreation*' to share similar activities to the gas networks. These are only included in our assessment because they were used by Ofgem and CEPA at RIIO-2.

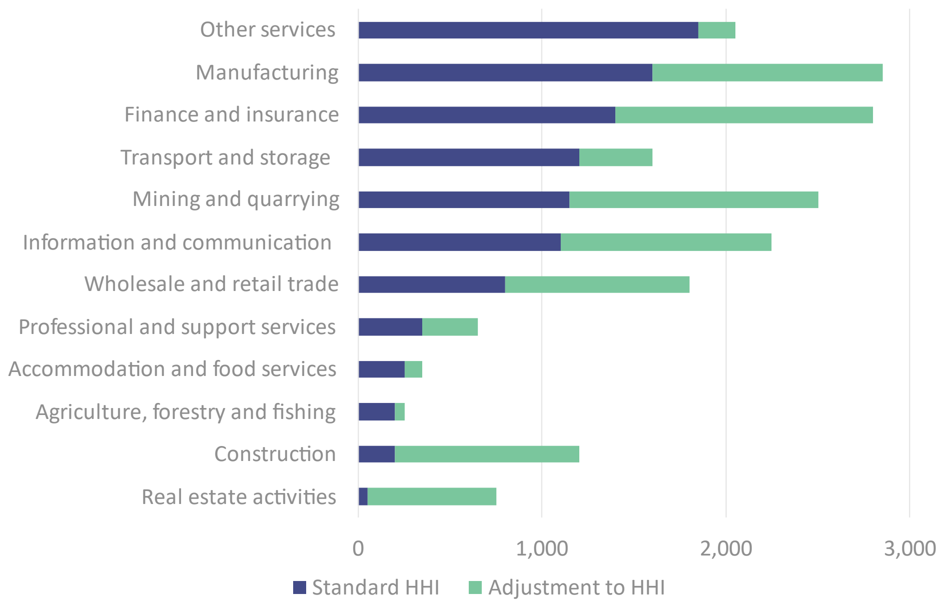
[40] *We do not consider this sensitivity (or any of our other sensitivities) appropriate for informing the OE challenge. This is because the sensitivities use assumptions that we do not recommend and are only included to test the robustness of our results.*

[41] *We note that this is not equivalent to the regulated sector (which only involves transmission and distribution networks), because it also contains a large number of different industries including energy generation and energy retail.*

Criterion 2: Competitiveness of industry

For Criterion 2, we assign "Amber" and "Green", using the Herfindahl-Hirschman Index (HHI) measure of market concentration, in the following way: (i) "Green" where the adjusted HHI is less than 1,000; and (ii) "Amber" where the adjusted HHI is greater than 1,000. The HHI informs the relative competitiveness of each industry[42] and is published by the CMA in its State of Competition report from April 2022.[43] Our threshold for "Green" and "Amber" is based on the CMA's explanation that *"[p]roduct markets with HHIs of more than 1,000 are generally considered to be concentrated, and those with HHIs of more than 2,000 to be highly concentrated"*.[44] Consistent with the CMA's 'highly concentrated' distinction (and as detailed in Annex 3), we also include a sensitivity[45] in which we remove all comparators that have an adjusted HHI that is greater than 2,000.[46] The figure below reports HHI and also the *adjusted* HHI figures. These adjusted figures account for the effect of common ownership and international trade, which the CMA considers affect competition, but are not included in the standard HHI measure.[47]

Figure 6: Standard and adjusted HHI across UK SIC sectors



Source: '[The State of UK Competition](#).' CMA (April 2022).
Notes: The sector names in the chart differ slightly to those used in the EU KLEMS database, e.g. (i). 'Transport and storage' in the above is equivalent to 'Transportation and storage' in EU KLEMS NACE II; (ii) 'Manufacturing' is equivalent to 'Total Manufacturing' in EU KLEMS NACE I; and (iii)

---

42 *The lower the HHI, the more competitive the industry. The bottom of the scale (i.e. perfectly competitive industry) is zero and the top of the scale (monopoly) is 10,000.*

43 *'The State of UK Competition.' CMA (April 2022); Figure 3.5; page 73.*

44 *'The State of UK Competition.' CMA (April 2022); para. 2.10.*

45 *We do not support the use of this, or any other sensitivity, to set the OE challenge for RIIO-3. This is because the sensitivities are based on assumptions that we do not recommend and are only included to test the robustness of our results.*

46 *We note that Figure 6 is based on the industries listed in the ONS dataset, rather than in the EU KLEMS dataset. In some cases, the ONS industries are more aggregated than those in the EU KLEMS dataset. For example, 'Transport and Storage' is broken down into further sectors in the EU KLEMS database, such as: (i) 'Land transport and transport via pipelines'; and (ii) 'Water transport'. Given the EU KLEMS data structure, we apply the Criterion 2 rating (as applied to the aggregated sector from the ONS data) to all of its constituent disaggregated sectors in the EU KLEMS dataset. For example, both 'Land transport and transport via pipelines' and 'Water transport' would be rated as "Amber" as 'Transport and storage' is rated as "Amber".*

47 *'The State of UK Competition.' CMA (April 2022); para. 3.1.*

'Professional and support services' corresponds to 'Professional, scientific and technical activities; administrative and support service activities' in the EU KLEMS database.

We assign any regulated or public sector industry as "Red". This is because they can be assumed to not operate in competitive markets, and so any TFP estimates may include a significant catch-up element, in addition to OE.

Criterion 3: Extent of scale effects

For Criterion 3, we calculate scale-related metrics across industries using company financial accounts aggregated by the FAME database. FAME contains detailed information on companies across the UK, and therefore allows us to determine which sectors are most similar to the regulated networks. This is because FAME can aggregate data across only the specific firms for which the OE challenge is calculated.[48] This provides an advantage over using EU KLEMS data, which would only allow comparisons to the much broader '*Electricity, gas, steam and air conditioning supply*' sector, rather than separating out the regulated networks. The metrics for these networks can then be compared to equivalent metrics aggregated across each potential comparator sector.

Criterion 3 is divided into three sub-criteria, which are as follows:

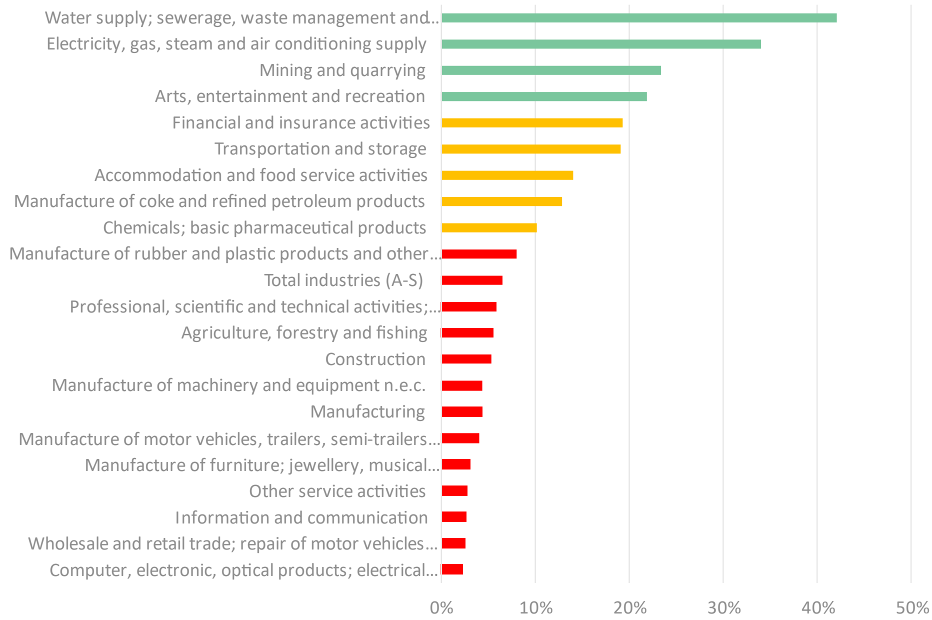
- **Criterion 3a: Fixed costs.** For this, we calculate the average fixed tangible asset-to-turnover ratio from 2013-2018[49] for each potential comparator sector and compare this to the gas networks. We note that the gas networks have a much higher ratio than *any* other sector. This is because the industry classifications are very broad; and thus contain a large variety of firm types; e.g., the '*Electricity, gas, steam and air conditioning supply'* sector contains a large number of retail firms, such as Octopus Energy (which will have a low level of capital intensity), as well as the regulated networks (which have a very high level). This means that, at a sectoral level, the average ratios will be less extreme than small groups of individual firms (such as gas networks, in this case). To assign "Red", "Amber", "Green" (RAG) ratings to each sector, we first sorted the sectors from highest to lowest in terms of their fixed tangible asset-to-turnover ratio. Comparator sectors that are *closest* to gas in terms of the proportion of costs that are fixed, will be those with the highest ratios (since gas networks have the highest ratio). We therefore assign the three-colour scale in the following way: (i) "Green" where the fixed tangible asset-to-turnover ratio is greater than 20%; (ii) "Amber" where the ratio is between 10% and 20%; and (iii) "Red" where the ratio is less than 10%. The figure below illustrates these ratios for each sector.

---

[48] *For each criterion, we compare the measure for each comparator to an average of all networks for which the OE challenge was jointly set for at RIIO-T2 and RIIO-GD2 (i.e. National Gas; Cadent; Northern Gas Networks, Scotland Gas Networks, Southern Gas Networks, Wales & West Utilities; and National Grid).*

[49] *This period was chosen because it: (i) omits the pandemic years to avoid any disruptions caused by COVID; and (ii) contains enough years of data to avoid annual variation influencing the results, while being recent enough to be relevant.*

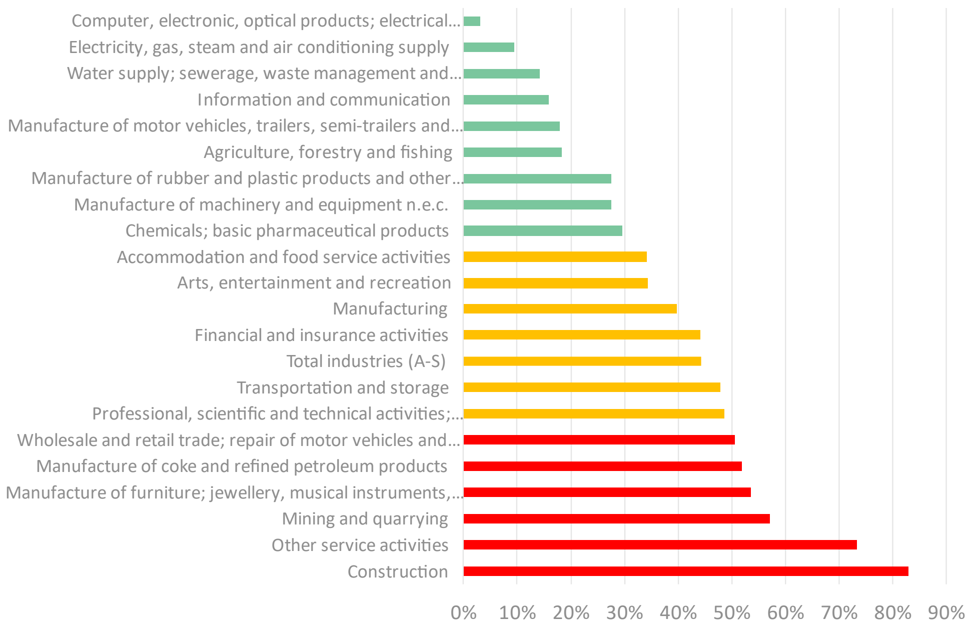
Figure 7: Fixed tangible asset-to-turnover ratio (2013-2018)



Source: Economic Insight analysis of FAME data

- **Criterion 3b: Fixed cost growth.** We calculate the growth rate of tangible fixed assets from 2013-2018 for each sector and compare it to the gas networks. We include this criterion on the basis that, where the growth rates are similar, this suggests that the comparator sector is similar to the networks in terms of its fixed cost accumulation over time. Following this, we then assigned the three-colour scale as follows: (i) "Green" where the absolute average divergence is less than 30%; (ii) "Amber" where the absolute average divergence is between 30% and 50%; and (iii) "Red" where the absolute average divergence is greater than 50%. The figure below shows the absolute divergence from gas networks.

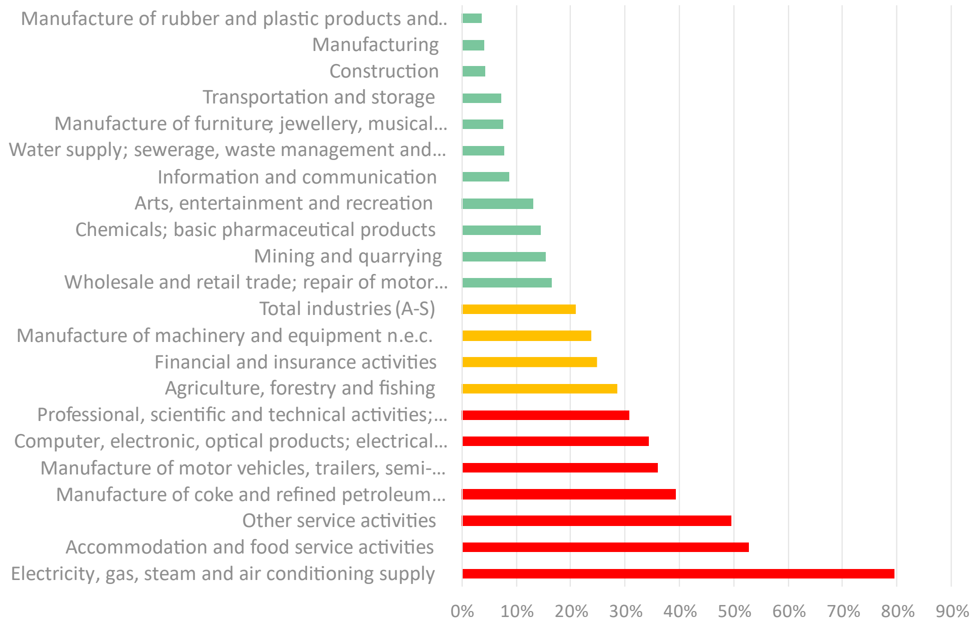
Figure 8: Absolute divergence in tangible fixed asset growth (2013-2018)



Source: Economic Insight analysis of FAME data

- **Criterion 3c: Turnover growth rate.** For this criterion, we calculate the growth rate of turnover in each year (this represents output growth) between 2013 and 2018; and then take the absolute difference between this rate and the rate for the gas networks. We include this criterion on the basis that, where the average divergence between growth rates is small, this suggests that scale effects are more similar over time. We then assign the following three-colour scale: (i) "Green" where the absolute average divergence is less than 20%; (ii) "Amber" where the absolute average divergence is between 20% and 30%; and (iii) "Red" where the absolute average divergence is greater than 30%. The figure below shows the absolute divergence from the networks for each possible comparator sector.

Figure 9: Absolute divergence in turnover growth rate (2013-2018)



Source: Economic Insight analysis of FAME data

Our choice of comparators

Notwithstanding our endeavour to ensure our selection of comparators is evidence-based, we recognise that no choice of comparators will ever be perfect. Therefore, although we have applied the above criteria, we recognise that there is an inherent level of subjectivity in comparator selection. Drawing the above together, we define our *'recommended set'* of comparators as those that fulfil all of the following three conditions across each of our criteria.

- For Criterion 1, the comparator must be defined as "Green" or "Amber", such that the activities being undertaken by firms working in the comparator industry are similar (at least in part) to the gas networks.

- For Criterion 2, the comparator must be defined as "Green" or "Amber", such that the industry is at least somewhat competitive.

- For Criterion 3, the comparator must be defined as "Green" in <u>at least one of</u> Criteria 3a, 3b and 3c, such that the magnitude and/or timing of scale effects are at least somewhat similar to the gas networks.

In addition, we also include '*Total industries*' in our *'preferred set'* of comparators (and the sensitivities detailed in Annex 3[50], unless specified otherwise). This reflects the inherent subjectivity in comparator choices, which means we think it is beneficial to include a metric that captures productivity changes across the entire UK 'on average' (i.e. not assuming that the gas networks would be either a 'low'- or a 'high'-productivity growth industry). Together, this yields the following comparator industries in our preferred set (more details of different groups of comparator industries and their performance against our assessment criteria can be found in Annex 2; and we compare our choices to CEPA's RIIO-2 approach in Annex 5):

- Total industries (A-S);
- Manufacturing;
- Chemicals; basic pharmaceutical products;
- Manufacture of rubber and plastic products and other non-metallic mineral products;
- Computer, electronic, optical products; electrical equipment;
- Manufacture of machinery and equipment n.e.c.;
- Manufacture of motor vehicles, trailers, semi-trailers and other transport equipment;
- Manufacture of furniture; jewellery, musical instruments, toys; repair and installation of machinery and equipment;
- Construction;
- Wholesale and retail trade; repair of motor vehicles and motorcycles; and
- Transportation and storage.

We also note that annual TFP growth rates for individual sectors can exhibit high variance (particularly problematic under the VA metric, as previously highlighted) and can be highly sensitive to changes in the data, whereas TFP growth for the economy as a whole is more stable. We illustrate this in the table below, which shows how the results of Ofgem's / CEPA's RIIO-2 benchmarking analysis change *when the underlying data is updated to the latest EU KLEMS release* (while using the same comparators and time period that were used by CEPA at RIIO-2).

[50] *We do not support using these sensitivities to inform the OE challenge because they are based on assumptions that we do not recommend. We only use the sensitivities to test the robustness of our recommended results and not to provide alternative estimates of OE.*

The table compares CEPA's results based on the EU KLEMS 2019 release, to our 'straight update'[51] of its analysis, using the EU KLEMS 2023 release. For CEPA's targeted comparator set, the range of productivity growth rates calculated from the old release is 0.2% to 0.8%. However, upon updating to the latest data, this range for the targeted comparator set undergoes a marked shift to -0.8% to -0.2%. In contrast (and crucially), the results for the economy-wide comparator set were much less sensitive to the updated data. The range for this comparator set changed from 0.4% to 1.0% before the update, to 0.4% to 0.8% after the update. We provide a full description of our update to CEPA's results in Annex 4.

Table 1: 'Straight update' of Ofgem's / CEPA's RIIO-2 benchmarking analysis

Measure	Expenditure category	Targeted comparator set		Economy-wide comparator set (weighted)	
		CEPA	EI update	CEPA	EI update
VA LP at constant K	opex	0.8%	**-0.8%**	1.0%	**0.5%**
VA TFP	capex, repex, opex	0.5%	**-0.6%**	0.9%	**0.8%**
GO LEMS at constant K	opex	0.3%	**-0.2%**	0.5%	**0.4%**
GO TFP	capex, repex, opex	0.2%	**-0.2%**	0.4%	**0.4%**

Source: CEPA's FD report at RIIO-2[52] and Economic Insight analysis of EU KLEMS data

Our method gives us a final preferred comparator set of 11 sectors (as listed above). This is a relatively large set of comparators. However, we consider this is appropriate (and important) to reduce the sensitivity of the results both to: (i) future updates to the data that are unrelated to changes in productivity performance; and (ii) the high variances in the annual TFP growth rates of individual sectors discussed above (which would be more problematic, were the VA metric used – and is notably problematic in the CEPA approach, as illustrated by our update of their prior assessment).

[51] *We use an identical method to that used for CEPA's RIIO-2 analysis, including using the same comparator set and time period (1997-2016). The only difference is that we use the latest release of EU KLEMS. We provide more detail in Annex 4. Note that this differs from the 'complete' update of CEPA's results, discussed previously.*

[52] *'RIIO-GD2 and T2: Cost Assessment –Advice on Frontier Shift policy for Final Determinations.' CEPA (November 2020); table 2.1*

OUR RECOMMENDED RANGE FOR OE AT RIIO-3 IS 0.2%-0.8% PA.

Results

To calculate the average TFP growth implied by a given set of comparators over a given time period, we: (i) first calculate the arithmetic mean[53] of annual TFP growth across the given time period for each of our comparator industries (using the EU KLEMS data); and then (ii) take a simple average across all industries in the comparator set of the annual TFP growth rates calculated in (i).

We think that it is highly likely that OE will be within this range at RIIO-3 (i.e. this is our recommended range). This is because it includes time periods that provide: (i) a lower bound given by a persistence of the UK's recent poor productivity growth performance; and (ii) a *likely* upper bound based on our view that the structural break in productivity growth is unlikely to completely unwind in the near-term. The range is based on: (a) our preferred comparator set; and (b) the time periods: 2010-2019, and a weighted average[54] of 1995-2019 and 1970-2007. The results for this range are shown in the table below.

Table 2: Results for our recommended range

Period	Average GO TFP growth (%)
Overall range	0.2% to 0.8%
2010-2019	0.2%
Weighted average of: (i) 1970-2007; and (ii) 1995-2019	0.8%

Source: Economic Insight analysis of EU KLEMS data

We provide further detail on the results for our preferred set in Annex 1 and present the results of our sensitivity analysis in Annex 3. We do not consider the results of our sensitivities appropriate for setting the OE challenge for RIIO-3 because they are based on assumptions and time periods that we do not recommend. The sensitivities are only included to test the robustness of our recommended results and not to provide alternative estimates for setting OE.

53 *We discuss the use of arithmetic means compared to geometric means in Annex 7 and explain why we believe arithmetic means are more appropriate.*

54 *As this allows for only partial unwinding of the UK productivity puzzle, at best. We provide extensive evidence on why a 'full unwinding' is highly unlikely in Chapter 3.*

3 The UK productivity puzzle

In this chapter we consider the evidence on whether the structural break in UK productivity growth (since 2008) is likely to limit the scope for gas networks to make productivity gains over RIIO-3. This topic has been subject to considerable debate in previous regulatory determinations. However, said debate has largely been informed by hypotheses and qualitative arguments, rather than a detailed consideration of evidence. This is undesirable, given the materiality of OE under price regulation. There are two key issues, which we address in turn in the following:

- Firstly, the structural nature of the flatline in productivity growth since 2008 and whether evidence suggests there is likely to be any material change (improvement) in productivity performance over the RIIO-3 time horizon.
- Secondly, whether regulated industries (such as gas networks) may be less (or more) affected by the factors driving the UK's flat and near-zero productivity growth.

3A. Structural break in UK productivity growth

UK productivity growth has flatlined since 2008

UK productivity growth has been flat since the 2008 financial crisis, as illustrated in Figure 10. This decline in productivity growth can be widely observed across most Western countries (although the persistence of low-to-stagnant productivity growth in the UK is particularly notable).[55] The academic literature contains no clear consensus as to why this has occurred; and so it has been termed: the '*productivity puzzle*'.

The *'productivity puzzle'* has been well documented in the literature,[56] and has been recognised by public bodies including the CMA. For example, the CMA's new research agenda includes investigating "*Innovation, investment, and productivity*" in relation to the UK's productivity puzzle, which it states is "*well documented, as is the degree to which investment and innovation activity has fallen since at least 2008*".[57]

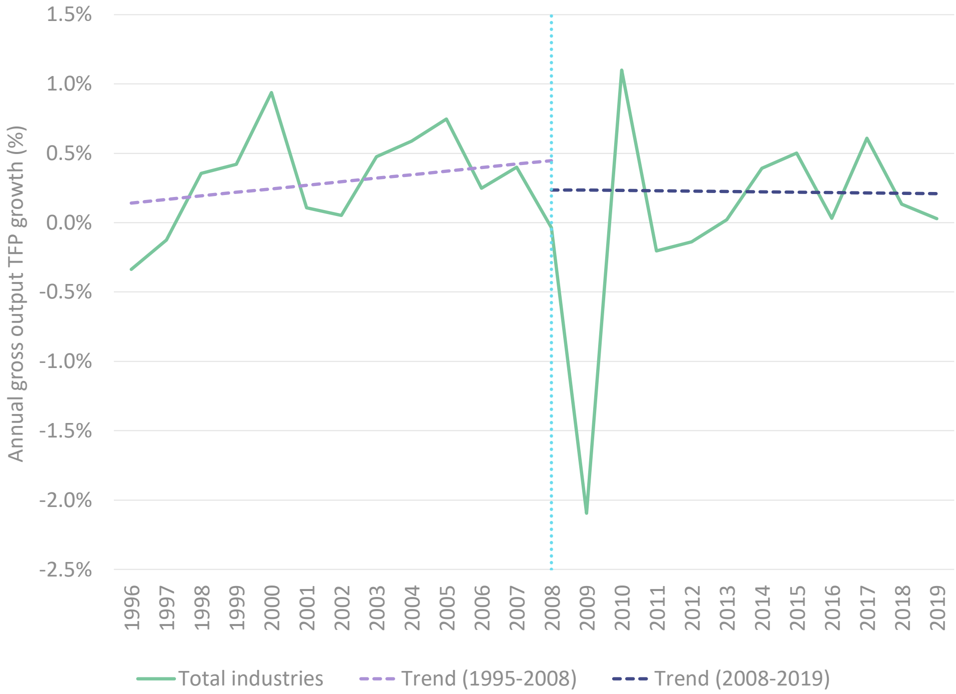
---

55 *For example, see: 'UK skills and productivity in an international context.' Aznar (2015); 'Solving the United Kingdom's productivity puzzle in a digital age.' Bughin et al. (2018); Below the aggregate: a sectoral account of the UK productivity puzzle.' Riley et al. (2018); and 'Productivity in the UK: Evidence Review.' The UK Productivity Commission (2022).*

56 *For example, see: 'The UK's Productivity Problem: Hub No Spokes.' Haldane, A. (2018); 'Below the aggregate: a sectoral account of the UK productivity puzzle.' Riley et al. (2018); 'Accounting for the UK Productivity Puzzle: A Decomposition and Predictions.' Goodridge et al. (2018); 'Is the UK productivity slowdown unprecedented?.' Crafts and Mills (2020); and 'Productivity in the UK: Evidence Review.' The UK Productivity Commission (2022).*

57 *'Strategy: Economic research.' CMA (March 2023)*

Figure 10: Average annual GO TFP growth rates with pre- and post-crisis trend lines



Source: Economic Insight analysis of EU KLEMS data

The productivity growth slowdown is unlikely to fully unwind over the RIIO-3 price control period

Setting aside the causes of the slowdown and their relevance to gas networks for now, a consideration that arises when setting OE is (as previously discussed) the choice of time periods. Seen through that lens, the relative weight one places on pre- / post-2008 data should depend on the extent to which we might expect those periods to be more / less reflective of achievable productivity growth over RIIO-3. In simple terms, one should consider whether the most up-to-date evidence suggests the productivity growth slowdown is likely to continue in the near-term or is likely to improve.

In this chapter, we therefore examine the following:

- The latest evidence on the UK's current economic outlook (forecasts and views from credible institutions). These suggest that an improvement in UK productivity growth is *unlikely* to occur in the near future.

- Views on future productivity performance potential, as provided by surveyed independent academic experts. The experts *do not* expect productivity growth to improve in the near future.

The UK's current economic outlook remains poor, which suggests that an improvement in UK productivity growth is unlikely to occur in the near future.

In its most recent Monetary Policy Report, the Bank of England forecasts that TFP growth will average 0.3% pa[58] from 2024 to 2026.[59] This suggests that the slowdown in productivity growth is likely to continue in the near future.

The UK economy's recent growth has also been poor. The latest ONS GDP data (released in February 2024) shows that the economy has failed to grow overall in the last two years (from Q1 2022 to Q4 2023), with real GDP falling slightly.[60] Additionally, the latest data also shows that at the end of 2023 the UK experienced a recession.[61] There were two consecutive falls in GDP on a quarterly basis, with a fall of 0.3% in Q4 2023 following a fall of 0.1% in Q3 2023.[62]

The latest short-term forecasts for the economy paint a similar picture to the above.[63] HM Treasury's latest consensus forecasts for the economy (which considered 19 new independent forecasts, all of which were made in February 2024, as well as the IMF's latest forecasts), suggest average forecast annual GDP growth of just 0.4% for 2024.[64] Similarly, the Bank of England has also recently forecast *"GDP growth to be broadly flat in Q4 and over coming quarters"*[65] (the strong correlation between GDP and productivity growth meaning these forecasts are highly consistent with the weak projections for productivity referenced above).

'The outlook for UK GDP growth is bleak for the foreseeable future' – NIESR

A number of credible institutions share this pessimistic outlook. For example, in November 2023 the National Institute of Economic and Social Research (NIESR) commented:

- *"The outlook for UK GDP growth is bleak for the foreseeable future. Although we do not expect to see a recession in the United Kingdom, we see growth of only 0.6 per cent this year and 0.5 per cent next year as the rapid tightening in monetary policy we saw between December 2021 and August of this year continues to bear down on output."*[66]

Similarly, in its October 2023 outlook for the UK economy, the Institute for Fiscal Studies (IFS) stated:

- *"The economic experience of the last three years is a harbinger of the kinds of supply shocks that are likely to come."*[67]

58 *This is calculated using an arithmetic mean of the forecast TFP growth across each the next three years (which are 2024, 2025 and 2026).*
59 *'Section 3 – In focus – The supply side of the economy.' Bank of England (February 2024).*
60 *'GDP in chained volume measures – real-time database (AMBI).' ONS (February 2024).*
61 *Note that a technical recession is defined as two consecutive quarters of contracting GDP.*
62 *'GDP first quarterly estimate, UK: October to December 2023.' ONS (February 2024).*
63 *'Forecasts for the UK economy: February 2024.' HM Treasury (February 2024).*
64 *'Forecasts for the UK economy: February 2024.' HM Treasury (February 2024); page 3.*
65 *'Monetary Policy Report.' Monetary Policy Committee (February 2024); page 27.*
66 *'The Outlook for the UK Economy.' NIESR (November 2023).*
67 *'UK Outlook: Fallout.' IFS (October 2023); page 3.*

- *"We expect weak margins and policy headwinds to drive a moderate recession through the first half of 2024. We expect GDP will fall 0.7% by next year, followed by growth of 0.4% in 2025."*[68]

Surveyed experts do not expect productivity growth to improve in the near future

For the purpose of a wider research exercise in relation to productivity in the UK, we have developed a working paper, exploring productivity drivers, including in relation to the post-2008 slowdown.[69] As part of that research, a survey was undertaken of the UK's leading independent academic experts in productivity analysis, including academics from all five UK productivity research centres.[70]

Below we highlight key results from the survey, both in relation to the prospects for future UK productivity growth over the next five calendar years (2024-2028), as well as the academics' expectations on which sectors will under or overperform UK productivity growth.

INDEPENDENT ACADEMIC EXPERTS IN PRODUCTIVITY EXPECT UK PRODUCTIVITY GROWTH TO BE LESS THAN 0.5% PA OVER THE NEXT FIVE CALENDER YEARS.

Most surveyed academic experts predict UK productivity growth will be less than 0.5% pa over the next five calendar years (which is significantly below the RII0-2 OE challenge of 1.0% pa[71]).

Evidence from the survey of academic experts suggests that *most do not expect material changes in UK productivity performance over the next five years (2024-2028)*, compared to the current very low productivity growth seen over the five most recent years (2016-2020).[72] Figure 11 below shows the relevant survey results.

---

68 *UK Outlook: Fallout.' IFS (October 2023); page 5.*

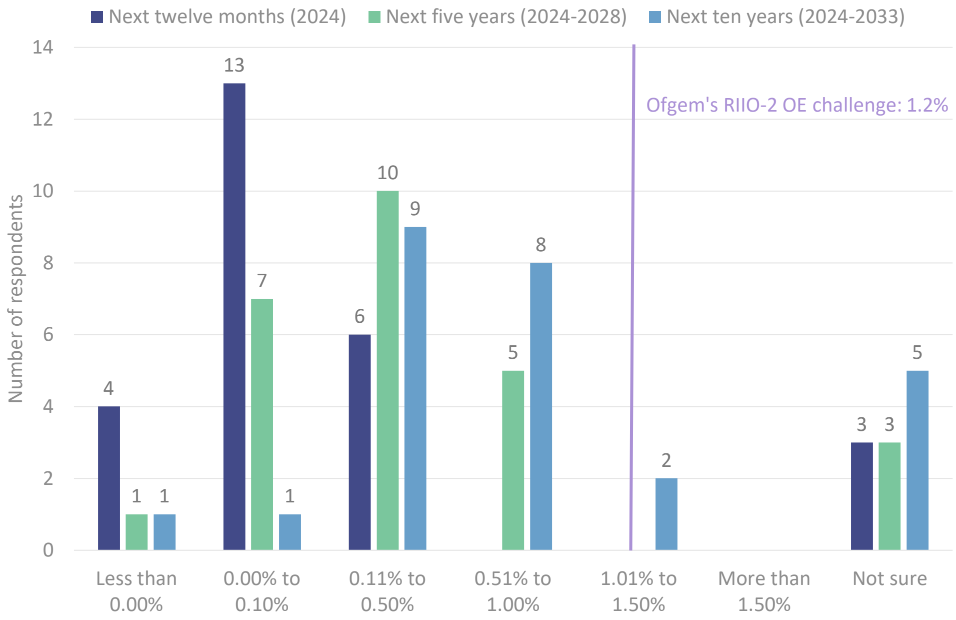
69 *The working paper has been submitted for publication and can be found here: 'The UK productivity puzzle: A survey of the literature and expert views.' Williams, S.; Glass, A.; Matos, M.; Elder, T.; and Arnett, D. (January 2024). The research was not funded by our clients, nor was the associated paper developed for the purpose of supporting regulatory submissions. Instead, it was an academic endeavour on the part of the listed authors. Participating academics engaged in the research of their own volition, and without any financial incentive.*

70 *These are: the Productivity Institute; Loughborough University's Centre for Productivity and Efficiency; the Programme on Innovation and Diffusion (POID); the Productivity Insights Network; and Lancaster University's Centre for Productivity and Performance.*

71 *The RIIO-2 OE challenge for gas networks was originally set at 1.2% by Ofgem, but this was reduced to 1.0% following CMA appeals. This is because the CMA removed the 0.2% innovation uplift included in the original challenge.*

72 *So defined as these were the five most recent years for which data was available, at the time the survey was run.*

Figure 11: Expectation of range of UK productivity growth (as measured by average annual % MFP, multifactor productivity) over the: (i) next twelve months (2024); (ii) next five calendar years (2024-2028); and (iii) next ten calendar years (2024-2033)



Source: Economic Insight survey of academic experts, N=26.

Over the next five calendar years (2024-2028):

- most academic experts (18 out of 23, or 78%, of those who provided an estimate) expect UK productivity growth to be 0.50% pa or below;
- 5 academic experts consider it will fall between 0.51% pa and 1.00% pa; whilst
- no academic experts expect productivity performance to be above 1.00% pa (and 3 academic experts did not provide an estimate / were unsure).

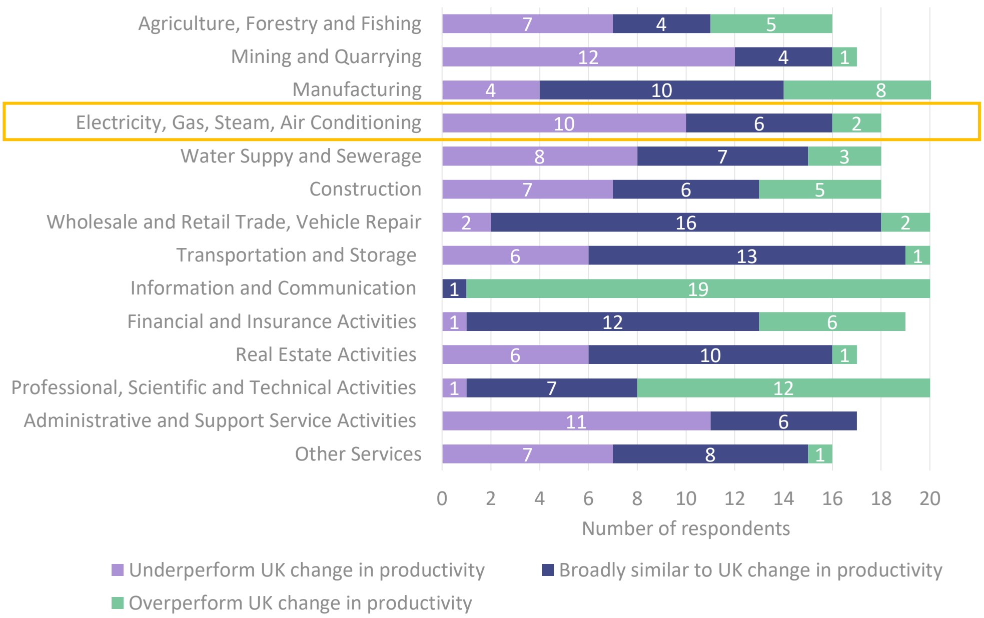
Of the 18 academic experts who expect productivity performance to be 0.50% pa or lower:

- 1 believes productivity performance will reduce over the next five calendar years (2024-2028), relative to its prevailing level of 0.00% pa over the most recent five calendar years for which data was available at the time of the survey (2016-2020).
- 7 believe productivity performance will remain broadly similar over the next five calendar years (2024-2028), relative to its prevailing level of 0.00% pa over the five most recent calendar years for which data was available (2016-2020). Reasons mentioned for this included that they consider there to be: no coherent infrastructure strategy; and no clear government policies.
- 10 respondents believe productivity performance will fractionally improve (i.e., be above zero, but no more than 0.50% pa) over the next five calendar years (2024-2028). Reasons cited for this include that they believe there will be increased private and public investment, along with increased stock (and quality) of human capital and increased openness to trade.

Moreover, leading academic experts expect energy sector specific productivity growth over the next five calendar years (2024-2028) to be lower than average UK productivity over the same time period.

As shown in Figure 11 (above) productivity growth over the next five calendar years (2024-2028) is expected to be 0.50% pa or lower by most independent academic experts. Those same experts were also asked how they expect productivity growth across each sector to change over the next five years, *compared to their expectations of how overall UK productivity might change over the next five years* (i.e., whether a sector might out- or underperform). Their responses are shown in Figure 12.

Figure 12: Expectations of how sector-level productivity might change over the next five years (2024-2028), compared to the expectation of how overall UK productivity might change relative to the last five years (2016-2020)



Source: Economic Insight survey of academic experts, N=23.

There are only 2 sectors for which the *majority* of academic experts expect outperformance relative to changes in UK-level productivity growth (‘*Information and Communication’; and ‘professional, scientific and technical activities’*). The academics identify a broader set of sectors (such as ‘*Wholesale and retail trade’*) where they expect the future productivity trend will remain broadly similar to that for the UK overall.

In relation to the ‘*Electricity, Gas, Steam and Air Conditioning Supply’* sector, 10 experts expect it to underperform the UK economy; with a further 6 expecting it to perform in line with the UK economy. That is, 16 experts (89% of respondents who provided an estimate for this question) expect the sector to perform below or in line with the UK economy, with most expecting it to perform below the UK economy. Only 2 expect it to outperform the UK economy.

3B. Regulated industries are broadly impacted by the same factors causing low productivity growth across the UK

IT IS REASONABLE TO CONSIDER WHETHER / HOW REGULATION MAY INFLUENCE PRODUCTIVITY. HOWEVER, PREVIOUS DISCUSSION OF THIS ISSUE HAS LACKED EVIDENCE.

Regulators have previously suggested that the factors causing low productivity growth in the UK (and most Western countries), post-2008, may not apply to regulated industries. However, the discussion around this has been largely informed by high-level reasoning rather than robust data and evidence. Moreover, without starting from a clear articulation of the factors driving the slowdown, there are inherent limits as to the degree one can reliably conclude regulation may have mitigated those factors. Nonetheless, it is important to recognise that it is entirely valid to consider what role, if any, regulation plays in influencing the achievable productivity of regulated companies.

In this section, we therefore address the above, whereby in turn we:

a. Firstly, identify and consider what the main factors causing the UK productivity growth slowdown are, based on evidence.

b. Secondly, consider the extent to which regulation might mitigate (or potentially increase) their impact on gas networks.[73]

What are the main factors causing the UK productivity growth slowdown?

To address this, we examine the following evidence: (i) academic literature, to identify the range of factors that affect productivity in general; and (ii) results from the survey of academic experts in productivity, to identify which factors are most important in causing the UK's productivity growth slowdown since 2008.

The literature identifies a range of factors affecting productivity

A working paper (2024)[74] we recently published, the scope of which included a review of the academic literature, identifies the key factors affecting UK productivity. Whilst there is no clear consensus as to which factors are *most relevant* to the current slowdown in the literature, it does identify the possible *range* of those factors. These include the following.

[73] *We note that regulation might also increase the impact of these factors on gas networks' specific productivity. That is to say, in the same way that is (in principle) plausible that regulation might assist productivity in some ways; it is similarly plausible that it might be detrimental to productivity in other ways. We do not consider the latter within the scope of this report but note that should the role of regulation prove to be the basis for some form of upwards adjustment to OE under Ofgem's determinations, it would be appropriate to revisit this.*

[74] *'The UK productivity puzzle: A survey of the literature and expert views.' Williams, S.; Glass, A.; Matos, M.; Elder, T.; and Arnett, D. (January 2024).*

- **Investment (private and public)** has long been associated with productivity growth.[75] A study by Syverson (2011) highlights the large body of research evidencing this.[76] Existing research covers the: (i) effect of private investment on productivity growth at both the *economy-wide, industry- and firm-level*; and (ii) influence of public investment on aggregate productivity growth. Additionally, the current productivity growth slowdown in the UK is strongly linked to underinvestment in several studies.[77] For example, Van Reenen (2023) states that the "*UK's productivity problem can be summed up in three words: investment, investment, investment. Or lack thereof*."[78] Thus, increasing private and public investment is frequently proposed as one of the key steps in remedying the UK's productivity puzzle.[79]

- **Infrastructure quality** is frequently associated with productivity growth in the literature. Infrastructure has been shown to positively affect productivity in both theoretical and empirical studies.[80] Relatedly, the poor state of the UK's infrastructure is identified in several papers as a key factor in explaining the flatlining of productivity growth since 2008.[81] One would expect this to adversely affect productivity quite broadly across the economy, as all sectors of the economy (including gas networks) rely on infrastructure as an input into production.

---

75 *See for example: 'The level of inventive activity.' Schmookler, J. (1954); 'The sources of measured productivity growth: United States agriculture, 1940-60.' Griliches, Z. (1963); 'Research expenditures, education, and the aggregate agricultural production function.' Griliches, Z. (1964); 'Sources of measured productivity change: Capital input.' Griliches, Z.; and Jorgenson, D. (1966); 'The explanation of productivity change.' Jorgenson, D.; and Griliches, Z. (1967); 'Productivity and the Role of Government.' Griffith, R.; and Simpson, H. (1998).*

76 *'What determines productivity?' Syverson, C. (2011).*

77 *See for example: 'Investing for prosperity: skills, infrastructure and innovation.' Besley, T.; Coelho, M.; and Van Reenen, J. (2013); 'Can intangible investment explain the UK productivity puzzle?' Goodridge, P.; Haskel, J.; and Wallis, G. (2013); 'Why should we care about productivity?' Pryce, V. (2015); 'Innovation, research and the UK's productivity crisis.' Jones, R. (2016); 'The UK's productivity puzzle: labour, investment and finance.' Chadha, J. (2017); 'What Is Holding Back UK Productivity? Lessons from Decades of Measurement.' Mason, J.; O'Mahony, M.; and Riley, R. (2018); 'Solving the United Kingdom's productivity puzzle in a digital age.' Bughin, J.; Dimson, J.; Hunt, V.; Allas, T.; Krishnan, M.; Mischke, J.; Chambers, L.; and Canal, M. (2018); 'A concerted effort to tackle the UK productivity puzzle.' Van Ark, B.; and Venables, A. (2020).*

78 *'Chronic under-investment has led to productivity slowdown in the UK.' Van Reenen, J. (November 2023).*

79 *See for example: 'Why should we care about productivity?' Pryce, V. (2015); 'Innovation, research and the UK's productivity crisis.' Jones, R. (2016); 'The Productivity Puzzle: It's the Lack of Investment, Stupid!' Herzog-Stein, A.; and Horn, G. (2018); 'The UK's productivity puzzle: labour, investment and finance.' Chadha, J. (2017); 'Productivity in the UK: Evidence Review.' UK Productivity Commission (2022).*

80 *See for example: 'The political economy of Leviathan.' Findlay, R.; and Wilson, J. (1987); 'Government, Trade, and Comparative Advantage.' Clarida, R.; and Findlay, R. (1992); 'International productivity differences, infrastructure, and comparative advantage.' Yeaple, S.; and Golub, D. (2007); 'Effects of Road Infrastructure on Employment, Productivity and Growth: An Empirical Analysis at Country Level.' Sotelsek, D.; and Laborda, L. (2019); 'Productivity impacts of infrastructure development in Asia.' Arif, U.; Javid, M.; and Khan, F. (2021).*

81 *See for example: 'Designing a new fiscal framework: Understanding and confronting uncertainty.' Chadha, J.; Küçük, H.; and Pabst, A. (2021); 'A concerted effort to tackle the UK productivity puzzle.' Van Ark, B.; and Venables, A. (2020); 'The UK's Productivity Problem: Hub No Spokes.' Haldane, A. (2018); 'Investing for prosperity: skills, infrastructure and innovation.' Besley, T.; Coelho, M.; and Van Reenen, J. (2013); 'Productivity in the UK: Evidence Review.' UK Productivity Commission (2022).*

- **Quality of the human capital stock** is an important factor determining productivity growth, as an educated and skilled workforce is known to raise productivity.[82] Views in the academic literature are somewhat mixed as to its relative importance *regarding the UK productivity puzzle specifically.* Although the UK appears to be close to the OECD average in terms of childhood education, primary and secondary schooling, there are clear and large disparities between socio-economic groups.[83] In addition, the UK is widely known to underperform in both further education and adult skills.[84] Additionally, there is evidence that, following both the 2008 financial crisis and the EU referendum, firms in the UK have been less likely to increase expenditure on worker training.[85] This is likely to lead to a worse skills mismatch in the future. For instance, research by the Industrial Strategy Council (2019) finds that 7 million additional workers (or 20% of the labour market) could be under-skilled for their job requirements by 2030.[86] Moreover, current evidence from the OECD (2023)[87] shows that, compared to other OECD countries, vocational training is less common in the UK.

- **Management quality** affects overall firm-level productivity growth. For example, one study identifies that, by coordinating the application of inputs, managers influence firm productivity.[88] Additionally, the literature frequently mentioned management quality as a contributing factor to the UK's slowdown in productivity growth since 2008.[89]

- The **misallocation of capital and labour** is commonly identified as a driver of productivity.[90] This is because, if inputs / resources are not put to their most productive use (at an economy-wide, industry or firm-level) then productivity will be harmed.

  – *Capital* misallocation is not so frequently referenced in relation to the UK's current productivity puzzle within the academic literature. However, a study by Pessoa and Van Reenen (2013) suggests that capital misallocation has increased in recent years and, therefore, has contributed to the slowdown.[91] They illustrate that: (i) the rate of bankruptcies and liquidations appears low; (ii) the cross-sectional variance of employment, output, and prices has increased across sectors; and (iii) there is an increased variance of productivity across firms within sectors.[92]

82 *'UK skills and productivity in an international context.' NIESR (2015).*
83 *'Putting together the pieces of the productivity puzzle: review article of productivity perspectives and productivity and the pandemic.' Van Ark, B. (2021).*
84 *'Putting together the pieces of the productivity puzzle: review article of productivity perspectives and productivity and the pandemic.' Van Ark, B. (2021).*
85 *'Firm investments in skills and capital in the UK services sector.' OECD (November 2020).*
86 *'UK Skills Mismatch in 2030.' Industrial Strategy Council (October 2019).*
87 *'Education at a Glance 2023: OECD Indicators.' OECD (2023).*
88 *'What determines productivity?' Syverson, C. (2011).*
89 *See for example: 'What Is Holding Back UK Productivity? Lessons from Decades of Measurement.' Mason, J.; O'Mahony, M.; and Riley, R. (2018); 'The UK's Productivity Problem: Hub No Spokes.' Haldane, A. (2018); 'Why is productivity slowing down?' Goldin, I.; Koutroumpis, P.; Lafond, F.; and Winkler, J. (2021); 'Putting together the pieces of the productivity puzzle: review article of productivity perspectives and productivity and the pandemic.' Van Ark, B. (2021).*
90 *'Misallocation and manufacturing TFP in China and India.' Hsieh, C.; and Klenow, P. (2009).*
91 *'The UK Productivity and Jobs Puzzle: Does the Answer Lie in Labour Market Flexibility?' Pessoa, J. P.; and Van Reenen, J. (2013).*
92 *'Micro-data: Perspectives on the UK Productivity Conundrum.' Field, S.; and Franklin, M. (2013).*

- In relation to *labour* allocation, evidence of its specific impact on the UK's productivity puzzle in the literature is mixed. While some studies consider that a poor allocation of labour resources is a contributing factor to the current slowdown in productivity growth,[93] others argue it is less important.[94]

- **Openness to trade** has been linked to improved productivity growth in theoretical studies.[95] However, it is not often considered a key explanatory factor for the UK's productivity puzzle.[96]

- **Government policy** has been shown to influence productivity growth. However, previous research on the contribution of government policy in the context of the UK's slowdown in productivity growth since 2008 is limited, and this is an area where further work is needed.[97]

- **Ownership structure of firms** has also been found to be a determinant of aggregate productivity.[98] Although there is *some* evidence that, generally, foreign-owned firms are more productive than domestically owned ones, it has not been frequently linked to the UK's productivity puzzle.[99]

Academic experts in productivity consider there to be five main drivers of the UK's current slowdown

In the survey of independent academic experts in productivity, respondents were asked to identify the factors that were most important in explaining the UK's slowdown, by:

- first, asking them to select which factors they consider explain the UK's productivity growth slowdown since 2008; and

- second, asking them to rank the five most important factors in explaining the slowdown.

[93] *See for example: 'Productivity: the route to Brexit success.' Dimson, J.; Hunt, V.; Mikkelsen, D.; Scanlan, J.; and Solyom, J. (2016); 'What Is Holding Back UK Productivity? Lessons from Decades of Measurement.' Mason, J.; O'Mahony, M.; and Riley, R. (2018); 'Putting together the pieces of the productivity puzzle: review article of productivity perspectives and productivity and the pandemic.' Van Ark, B. (2021); 'Productivity in the UK: Evidence Review.' UK Productivity Commission (2022).*

[94] *'Accounting for the UK productivity puzzle: a decomposition and predictions.' Goodridge, P.; Haskel, J.; and Wallis, G. (2018).*

[95] *'The impact of trade on intra-industry reallocations and aggregate industry productivity.' Melitz, M. (2003); 'What determines productivity?' Syverson, C. (2011).*

[96] *'The UK's Productivity Problem: Hub No Spokes.' Haldane, A. (2018).*

[97] *'The Politics of Productivity: institutions, governance and policy: Working Paper No. 015.' Pabst, A.; and Westwood, A. (2021).*

[98] *'Trade, location of economic activity and the MNE: A search for an eclectic approach.' Dunning, J. H. (1977); 'Toward an eclectic theory of international production: Some empirical tests.' Dunning, J. H. (1980); 'The eclectic paradigm of international production: A restatement and some possible extensions.' Dunning, J. H. (1988).*

[99] *'The UK's Productivity Problem: Hub No Spokes.' Haldane, A. (2018); 'Productivity in the UK: Evidence Review – First report of the UK Productivity Commission.' UK Productivity Commission (2022).*

Surveyed academic experts consider that the five most important factors explaining the productivity growth slowdown are the: (i) extent of **private investment**; (ii) **quality of infrastructure**; (iii) extent of **public investment**; (iv) quality of the **human capital stock**; and (v) quality of **firm management.**[100]

Table 3 below summarises the key results.

Table 3: Factors driving lower UK MFP growth since 2008, ranked by most important

Ranked ... important factor by expert ➔ Factor 🡻	most	second most	third most	fourth most	fifth most	Total mentions as one of the five most important factors	Total mentions as factor explaining UK MFP growth
Private investment	11	1	4	0	1	**17**	**17**
Quality of infrastructure	1	5	2	6	1	**15**	**17**
Public investment	1	6	2	3	2	**14**	**15**
Human capital stock	4	2	2	1	2	**11**	**14**
Firm management quality	2	4	2	1	2	**11**	**13**
Capital allocation across industries	2	3	1	1	2	**9**	**9**
Openness to trade	0	1	1	4	1	**7**	**7**
Labour allocation across industries	3	0	2	1	0	**6**	**8**
Regulatory and competition policy	0	2	2	0	0	**4**	**4**
Other factors	0	1	2	0	0	**3**	**3**
Government fiscal policy	1	0	0	1	0	**2**	**3**
Mix of firm ownership structures	0	0	2	0	0	**2**	**3**
Government monetary policy	0	0	0	0	0	**0**	**0**

Source: Economic Insight survey of academic experts, N=26.

[100] *Please note that for the subsequent analysis we refer to investment as encompassing both private and public investment.*

To what extent might regulation mitigate the impact of these factors on gas networks?

WE ASSESS A RANGE OF EVIDENCE TO MORE PRECISELY IDENTIFY THE CAUSAL FACTORS OF THE UK PRODUCTIVITY GROWTH SLOWDOWN. THIS INCLUDES A LITERATURE REVIEW, RESULTS FROM THE SURVEY OF INDEPENDENT ACADEMIC EXPERTS, AND TRENDS IN INVESTMENT LEVELS IN THE UK AND THE ENERGY SECTOR.

To address this question, having more precisely identified the causal factors of the UK productivity growth slowdown above, we assess a range of evidence, including: (i) academic and practitioner literature; (ii) results from the survey of independent academic experts; and (iii) trends in investment levels and growth in the UK and for the energy sector. The remainder of this section is structured as follows:

- First, we outline the ways in which it has previously been proposed that regulation might mitigate the impact of the productivity growth slowdown on regulated industries.
- Second, we evaluate the possible impact of regulation, discussing each of the key productivity drivers identified previously.
- Third, we examine drivers of across-sector variation in productivity – and what that implies for the impact of regulation – especially as regards regulatory innovation funding and incentives.
- Finally, we summarise our conclusions on the potential for regulation to mitigate the impact of the slowdown on gas networks.

Suggested ways in which regulation might mitigate the impact of the productivity growth slowdown on regulated industries

Sectoral regulators and the CMA have previously suggested that regulated industries are less affected by the UK productivity growth slowdown since 2008. Reasons put forward for this are:

- In giving evidence to the CMA during the PR19 redeterminations, Ofwat highlighted reasons forwarded by its consultants; primarily being that water companies are shielded from demand reductions and reductions in investment.[101]

[101] *'Additional Evidence on Some Points Relating to Frontier Shift.' Europe Economics (2020); page 16.*

- The CMA agreed with the argument that water companies may be less impacted (than other sectors) by factors causing the productivity flatline in the PR19 redeterminations. The CMA noted that the water sector might be less affected by reduced investment, due to the certainty provided by the regulatory regime, stating that *"[t]here were reasons which indicated that water companies were likely to be less impacted than other sectors. For example, the water sector would be less impacted by lower capital investment given the certainty provided by the regulatory regime and the innovation fund encouraging investments in new technologies."*[102] However, the CMA's overall conclusions on OE at that time were, in part, also informed by it believing that UK productivity performance was likely to improve over the following 5 years (which, so far, it has not).[103]

- Ofgem's consultants suggested *"[p]lacing less weight on the wider productivity slowdown in recent years, which would effectively see the productivity puzzle as being less relevant for regulated utility sectors – e.g. because of greater revenue and investment certainty in the regulated sectors"*[104], during the RIIO-2 price control period. They also proposed Ofgem ought to consider "*the benefits of innovation funding provided in RIIO-1 in improving the potential for the network companies to achieve productivity levels closer to those in the better performing competitive sectors*"[105] when setting the OE challenge.

- In its Final Determinations, Ofgem stated that: "*the innovation funding provided by consumers since 2007 should deliver efficiency benefits over and above those achieved in the wider economy, in comparator sectors, and beyond the range indicated by EU KLEMS."*[106] In the RIIO-2 energy appeals, the CMA found that GEMA (Ofgem) had not erred in double-counting innovation funding in the OE challenge.[107] However, the CMA did find that GEMA had made errors in aspects of its decision to set the innovation uplift at 0.2% (and so, overturned this).[108]

102 *'Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations: Final report.' CMA (March 2021); para. 4.537.*
103 *'Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations: Final report.' CMA (March 2021); para. 4.537.*
104 *'RIIO-GD2 and T2: Cost Assessment – Advice on Frontier Shift policy for Final Determinations.' CEPA (November 2020); page 8.*
105 *'RIIO-GD2 and T2: Cost Assessment – Advice on Frontier Shift policy for Final Determinations.' CEPA (November 2020); page 8.*
106 *'RIIO-2 Final Determinations – Core Document (REVISED).' Ofgem (February 2021); para. 5.26*
107 *Cadent Gas Limited, National Grid Electricity Transmission plc, National Grid Gas plc, Northern Gas Networks Limited, Scottish Hydro Electric Transmission plc, Southern Gas Networks plc and Scotland Gas Networks plc, SP Transmission plc, Wales & West Utilities Limited vs the Gas and Electricity Markets Authority: Final determination Volume 2B: Joined Grounds B, C and D.' CMA (October 2021); para. 7.412.*
108 *Cadent Gas Limited, National Grid Electricity Transmission plc, National Grid Gas plc, Northern Gas Networks Limited, Scottish Hydro Electric Transmission plc, Southern Gas Networks plc and Scotland Gas Networks plc, SP Transmission plc, Wales & West Utilities Limited vs the Gas and Electricity Markets Authority: Final determination Volume 2B: Joined Grounds B, C and D.' CMA (October 2021); para. 7.802.*

- In hearing the RIIO-2 energy appeals, the CMA, found that: "*there are reasons why the energy companies may be less impacted than other sectors. For example, the comparative certainty provided by the regulatory regime could facilitate investment.*"[109] However, it also noted that there was a risk that attaching undue weight to the post-crisis period risked setting an OE challenge that was too high: "*we agree with the appellants that an approach which placed insufficient weight on the lower productivity since 2008 could lead to an overestimate of the appropriate OE challenge*."[110] This latter observation is important in the context of the CMA choosing not to overturn GEMA's weighting of the pre-post crisis time periods. In this case specifically, the CMA was considering the narrow question of whether GEMA was 'wrong' rather than what the most appropriate method was 'from scratch.'

In summary, sectoral regulators and the CMA have advanced various reasons as to *why* regulation might mitigate the impact of the UK productivity growth slowdown on regulated companies' productivity. Overall, the two key reasons mentioned are that: (i) regulation makes investment more secure, the implication being that regulated sectors may be less exposed to underinvestment; and (ii) regulators have introduced specific incentives to boost innovation.

We agree that the above are sound *in principle* considerations – and it is reasonable to examine whether, and how, regulation might affect productivity. However, in order to more robustly determine the impact of regulation *in practice,* it is essential to consider its effects in relation to specific productivity drivers, as identified using evidence. In the following, we therefore address this.

The extent to which regulation may affect the most important factors determining the UK's productivity growth slowdown

Here, we explore the extent to which regulation may mitigate the impact (on gas networks) of the following factors identified in the evidence as determining the UK's productivity growth slowdown: (i) investment; (ii) infrastructure quality; (iii) quality of human capital stock; and (iv) management quality. Evidence is provided for each, in turn.

Investment

Investment, as previously mentioned, is a key factor in determining productivity growth more generally. Furthermore, the available evidence also suggests underinvestment helps explain the productivity growth slowdown in the UK since 2008. Whilst various factors affect investment, fundamentally it comes down to whether investor returns are attractive relative to the risks they face.

---

[109] *'Cadent Gas Limited, National Grid Electricity Transmission plc, National Grid Gas plc, Northern Gas Networks Limited, Scottish Hydro Electric Transmission plc, Southern Gas Networks plc and Scotland Gas Networks plc, SP Transmission plc, Wales & West Utilities Limited vs the Gas and Electricity Markets Authority: Final determination Volume 2B: Joined Grounds B, C and D.' CMA (October 2021); para. 7.87.*

[110] *'Cadent Gas Limited, National Grid Electricity Transmission plc, National Grid Gas plc, Northern Gas Networks Limited, Scottish Hydro Electric Transmission plc, Southern Gas Networks plc and Scotland Gas Networks plc, SP Transmission plc, Wales & West Utilities Limited vs the Gas and Electricity Markets Authority: Final determination Volume 2B: Joined Grounds B, C and D.' CMA (October 2021); para. 7.80.*

Some of the usually uncertain and unknown risk-return balance is less uncertain through the existing regulatory framework, which provides investors with a minimum return (or a narrower range of returns, relative to unregulated sectors). Therefore, in principle, through this channel, regulation *may* mitigate the prospects of underinvestment harming productivity performance for gas networks, relative to the wider economy.

'The above [key drivers of investment] *is dependent on Ofgem setting price controls such that investment in the sector is attractive, relative to its risk (i.e., setting the 'right' amount of allowed revenues).'*

In practice, however, the above depends on Ofgem setting price controls so that investment in the sector is attractive, relative to its risk (i.e., setting the 'right' amount of allowed revenues). In other words, if price controls are set 'too tight' (which includes setting the OE challenge 'too high'), regulation would actually have the opposite effect, leading to investment being below the level necessary to give customers and society their desired outcomes (and vice-versa). In addition, one also needs to consider how investment in other sectors affects gas networks' performance. For example, as we show later, where the UK *in general* is underinvesting in skills / human capital / other inputs used by gas networks, this will impact their productivity, even where regulation (to some degree) mitigates investment *in the gas sector itself.*

To understand *whether* regulation does, in fact, mitigate the impact of investment on productivity growth for gas networks, in the following we consider: (i) trends in energy sector investment relative to overall UK investment;[111] and (ii) implications of the extent of underinvestment across the UK overall.

<u>Trends in energy sector investment relative to overall UK investment</u>

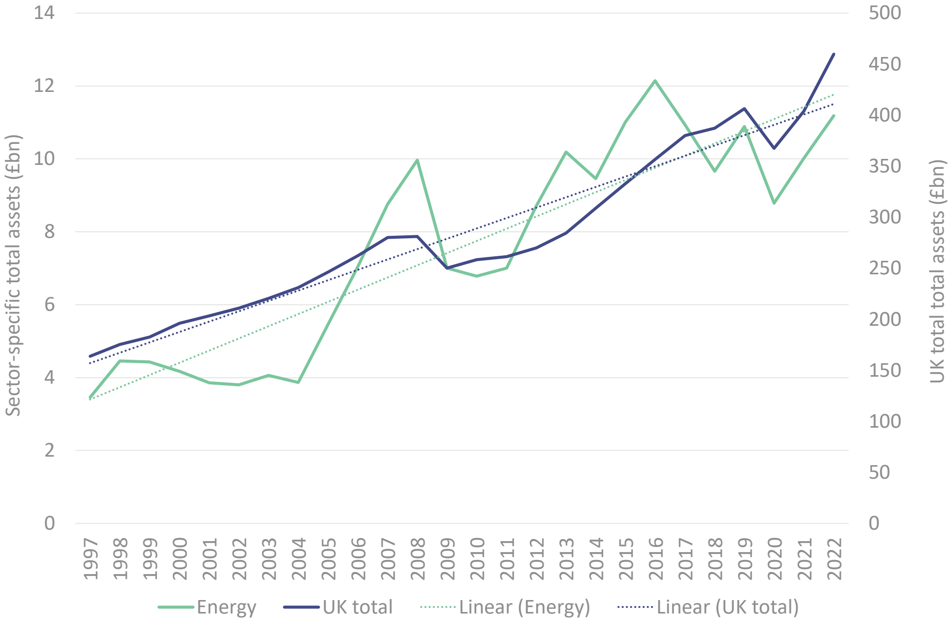
UK INVESTMENT TRENDS IN THE UK ARE MIRRORED IN THE ENERGY SECTOR.

We can see that total investment (gross fixed capital formation - GFCF) in the energy sector broadly follows a similar trend to the UK from 2006 onwards, albeit being more volatile than the UK overall (see Figure 13).[112] This similar trend suggests that, relative to the UK, the energy sector has <u>not</u> experienced systematically more, or less, growth in investment over time. Accordingly, if it is considered that the UK overall has been underinvesting since 2008 (and that this has contributed to the productivity growth slowdown – which seems to be widely accepted), this suggests that this should also be the case for the energy industry.

[111] *Note, as mentioned previously, we are not able to more granularly consider investment of just gas networks compared to other sectors in the UK economy using ONS and OECD data.*

[112] *Note that the 'Energy' investment data in Figure 13 and Figure 14 refers to 'Electricity, gas, steam and air conditioning supply.'*

Figure 13: Annual gross fixed capital formation, total assets (£bn, current prices)



Source: Economic Insight analysis of ONS GFCF (investment) by asset and industry data

Figure 14, which shows trends in total assets for the energy sector and the UK, further shows that the energy sector has followed a similar trend in investment to the UK since 2007. Crucially, the overall growth in investment has actually been *lower* for the energy industry than for the UK since 2007.

Figure 14: Percentage change in annual gross fixed capital formation (total assets) from 2007

Source: Economic Insight analysis of ONS GFCF (investment) by asset and industry data

*Notes: **light blue dotted lines** illustrate start of (i) TPCR4 and GDPCR1[113]; (ii) RIIO-GD1 and RIIO-T1; and (iii) RIIO-GD2 and RIIO-T2.*

[113] These two periods start one year apart with TPCR4 beginning in 2007 and GDPRC1 beginning in 2008. However, for presentational purposes, we simplify the diagram by using a single line to indicate the start of both periods.

In addition, we can see that:

- in the TPCR4 (2007-2013[114]) and GDPCR1 (2008-2013) periods, over 2009 to 2012 GFCF investment fell proportionally *more* in the energy sector than for the UK overall.

- in the RIIO-T1 and RIIO-GD1 period (2013-2021), energy GFCF growth followed a *similar* trend to the UK overall from 2013 to 2016. However, over 2017 and 2018, energy GFCF investment decreased, whilst it increased for the UK overall.

- from 2019 onwards, energy GFCF growth followed a similar trend to the UK overall; however, at a *lower* level, relative to 2007.

- finally, during both the 2008 financial crisis and the COVID-19 pandemic, energy GFCF growth dropped at a *faster* rate than for the UK overall.

Therefore, it appears that, compared to the UK economy as a whole, energy sector investment broadly follows a similar pattern, but is relatively lower in recent times. Moreover, in the energy sector reductions in investment at times of downturns actually appear larger than for the UK overall. In summary, when one examines the data, there is limited evidence to indicate that regulation is materially affecting investment in the energy industry, relative to the (likely more important) drivers of investment attractiveness across the economy as a whole.

The above is an unremarkable result, as the literature identifies a range of different types of shocks that have affected the UK economy, reducing investment certainty and harming the overall investment environment for businesses and investors. For example, relevant events encompass: (i) the 2008 financial crisis; followed by (ii) the 2010s austerity in the UK; (iii) the UK's withdrawal from the European Union in 2020; (iv) the COVID-19 pandemic in 2020; and now (v) global conflicts.[115] Even if the regulatory regime has some impact on investment-returns certainty, these external shocks still affect regulated companies' ability to attract investment. Most pertinently, it is likely determined by the overall UK investment environment.

<u>Evidence suggests there is economy-wide underinvestment in the UK</u>

There is wide agreement in the literature that the: (i) underinvestment in the UK is chronic, i.e., it is not a recent problem; and (ii) weak investment is broad-based, across many industries.[116] For example, it is extensively recognised that, in terms of investment, the UK persistently lags behind other comparable countries.[117] As illustrated in Figure 15, since 1997, UK total investment (GFCF) as a share of GDP has consistently lagged G7 countries.

---

[114] *Although TPCR4 expired in 2012, the roll over period expired in 2013. The reported range here therefore refers to the end of the roll over period (or, equivalently, the start of RIIO-T1).*

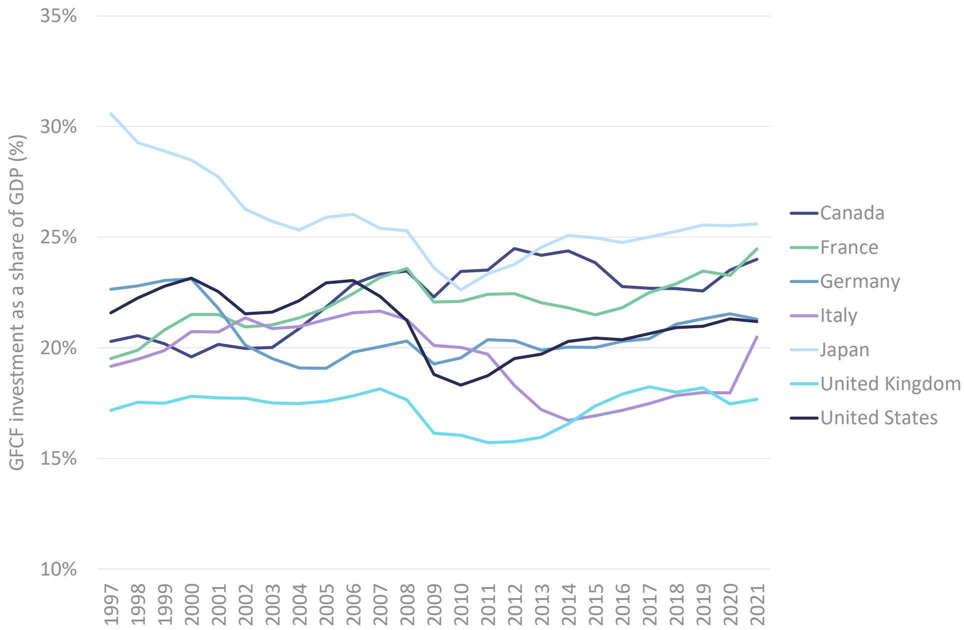
[115] *'Business investment: Not just one big problem.' Institute for Government (August 2022); 'The Productivity Agenda.' The Productivity Institute (2023).*

[116] *'The Productivity Agenda.' The Productivity Institute (2023); page 9.*

[117] *'Investment in the UK: Longer term trends.' Bennett Institute for Public Policy (November 2023).*

STRUCTURAL FEATURES OF THE UK ECONOMY HELP EXPLAIN THIS PERSISTENT UNDERINVESTMENT.

Figure 15: Total investment (GFCF) as a share of GDP



Source: Economic Insight analysis of OECD data

Various organisations suggest that structural features of the UK economy explain this persistent underinvestment, such as, for instance, the UK's business culture and its institutions, leading to short-termism and aversion to investment.[118] Where these factors cause the underinvestment, they are economy-wide and so are unlikely to be mitigated by regulation (i.e., a business culture is a UK-wide phenomenon).

Quality of infrastructure

As previously explained, infrastructure is a key factor in determining productivity growth in general and also helps explain the productivity growth slowdown since 2008.[119]

Evidence that the UK has low quality of infrastructure abounds:

- A report by the OECD (2015) observed that the UK's public spending on infrastructure has been lower over the past 30 years, compared to other OECD countries.[120] It also found that the perceived quality of UK infrastructure assets is *lower than in other G7 countries*, although it is close to the OECD average. The report found:
  - Investment in UK roads was considerably below the level in Germany and France, and the perceived quality of the UK's **road system** was worse than in most OECD countries of similar size and wealth.[121]

118 'Britain's Investment Gap.' TUC (2014); 'Britain's productivity problem is long-standing and getting worse.' The Economist (June 2022); 'Business investment: Not just one big problem.' Institute for Government (August 2022).

119 'Infrastructure policies and investment.' UK Parliament (March 2021).

120 'Improving infrastructure in the United Kingdom: Economics Department Working Papers No. 1244.' OECD (July 2015).

121 'Improving infrastructure in the United Kingdom: Economics Department Working Papers No. 1244.' OECD (July 2015); page 15.

- The **railway sector** in the UK was considered much improved over the last 20 years by the OECD. However, significant concerns around overcrowding remained.[122] In 2021, the UK government additionally noted that the railway's "*performance was disappointing and passengers' biggest priority for improvement was punctuality*."[123]
- The UK is significantly capacity constrained in relation to **air transport**.
- Further investment is needed to ensure adequate capacity at increasingly congested UK **sea ports.**[124] Notwithstanding this, the report found that the quality of the UK port infrastructure was perceived as relatively good, but not as high as in the best-performing OECD countries.

- Looking at the UK **healthcare infrastructure**, a report by the King's Fund (2023) found that the "*NHS, which sits at the core of the UK health system, is neither a leader nor a laggard when compared to the health systems of 18 similar countries*."[125] Further, it found that compared to other countries the UK's health spending per person was below average; and that the UK lagged behind other countries in its capital investment. This resulted in fewer key physical resources compared to many of its peers, such as computerised tomography (CT) and magnetic resonance imaging (MRI) scanners, and hospital beds.
- The National Infrastructure Commission (NIC) observed that "*much of the country's infrastructure is under strain, not keeping pace with population growth and modern requirements.*"[126]
- The NIC (2023) further noted that "*there are significant deficiencies that are holding the UK back. There has been under investment in transport systems in regional English cities, no major water resource reservoirs have been built in England in the last 30 years, too many properties are at risk of flooding, and recycling rates have not increased in a decade.*"[127]

All companies and sectors depend on infrastructure quality in the production of goods and services. This is what makes it so important to productivity performance. Whilst regulation of gas networks might (if it mitigated underinvestment – which we addressed above) result in higher quality *gas network* infrastructure, it does not (and cannot) protect gas networks from the effect of low-quality infrastructure *across the UK economy (e.g., low-quality transport links).* Thus, there are no grounds to suppose regulation can mitigate the productivity growth slowdown on gas networks companies based on this factor.

---

122 *'Improving infrastructure in the United Kingdom: Economics Department Working Papers No. 1244.' OECD (July 2015); page 17.*
123 *'Great British Railways: The Williams-Shapps Plan for Rail.' Department for Transport (May 2021); page 13.*
124 *'Improving infrastructure in the United Kingdom: Economics Department Working Papers No. 1244.' OECD (July 2015); page 22.*
125 *'How does the NHS compare to the health care systems of other countries?' The King's Fund (June 2023).*
126 *'Congestion, Capacity, Carbon: Priorities for national infrastructure: consultation ahead of National Infrastructure Assessment.' National Infrastructure Commission (October 2017); page 2.*
127 *'The Second National Infrastructure Assessment.' National Infrastructure Commission (October 2023); page 8.*

Quality of human capital stock

A key factor driving productivity in general (and explaining the slowdown in growth since 2008) is the quality of human capital stock. The low quality of the UK's human capital stock is evidenced by the following:

- The Department for Business, Innovation and Skills (BIS) (2015) found that, although the UK performs relatively well in terms of higher skills (bachelor's degree and above) compared to other countries, its comparative performance in intermediate (practical, technical, and occupational) skills was a concern.[128]

- In addition, the UK Commission for Employment and Skills (2015) showed that the UK ranks below average compared to the OECD and EU averages on both lower (e.g., below upper secondary) and intermediate skills (e.g., upper secondary), whereas it scores above average for higher skills (e.g., tertiary).[129]

- Multiple shortcomings in relation to the UK's skills were highlighted in the CIPD's (2017) response to the government's Industrial Strategy Green Paper.[130] These included: (i) for literacy and numeracy among 16-24 year olds, England and Northern Ireland together ranking in the bottom four OECD countries; (ii) for young people's computer problem-solving skills, the UK ranking bottom (out of 19 countries); (iii) UK employers spending less on training than other major EU economies, and less than the EU average; (iv) for job-related adult learning participation, the UK ranking fourth from the bottom on the EU league table.

Similarly to (physical) infrastructure set out previously, all companies and sectors will be affected by the quality of human capital stock available across the UK economy. This is because all sectors of the economy are (in the broadest sense) drawing from the same pool of applicants / labour market. The human capital stock (e.g., knowledge and skills) within the labour market is not, in any way, influenced by the regulation of gas networks. Thus, there is no evidence to suggest that regulation can mitigate the impact of this important driver of the UK productivity growth slowdown on gas networks.

Management quality

Another important determinant of productivity growth, and a driver of the slowdown since 2008, is firm management quality. There is a range of evidence in relation to the UK's shortcomings in this regard:

- BIS (2012) finds that a deficit in management and leadership skills is reported by nearly three quarters of organisations in England.[131] Findings include UK managers: (i) being less qualified; (ii) being under-trained; (iii) lacking key skills; and (iv) not strategically applying leadership and management skills.

---

128 *'UK skills and productivity in an international context.' BIS (December 2015).*
129 *'UK Skills Levels and International Competitiveness 2014.' UK Commission for Employment and Skills (October 2015).*
130 *'From 'inadequate' to 'outstanding': making the UK's skills system world class.' CIPD (April 2017).*
131 *'Leadership & management in the UK – The key to sustainable growth.' BIS (July 2012); page 15.*

- The UK's deficit in management quality compared to the US, Germany, Japan, and Sweden is highlighted by Bloom and Van Reenen (2007).[132] They also find that the cause of the productivity gap to those countries is likely to be this management deficit.

- Compared to other countries, managers in the UK tend to have lower levels of formal education. The share of managers with at least a tertiary education in the UK was below the EU average in 2017, and well below the best performers.[133]

Companies across all sectors in the UK will be broadly drawing from the same pool of managers, similarly to the quality of the human capital stock issue set out previously. Therefore, there are limited reasons to consider that regulation can mitigate the impact of management quality on gas networks specific productivity.

Understanding drivers of across-sector variation in productivity

Whilst the productivity growth slowdown appears pervasive across many Western countries, and most sectors of the UK economy, variation across industries is observable. This variation can provide further insights as to whether, and to what extent, regulation might mitigate the effects of the slowdown on regulated companies. Accordingly, in the following, we discuss:

– what factors most drive sectoral variation in productivity growth;

– how the main factor driving across-sector variation is the 'scope for technological change' and how this is intrinsic to industry characteristics; and

– the implications of this for any impacts of regulation on productivity.

The 'scope for technological' change is identified as the most important driver of variation in sectoral productivity growth

Within the academic experts survey, respondents were asked:

– which factors they thought explained variations in productivity growth across sectors; and

– of those factors, which they considered to be the five most important.

The most important factor identified as explaining variations in productivity growth across sectors is the **scope for technological change**. Table 4 shows the results.

---

132 'Measuring and Explaining Management Practices Across Firms and Countries.' Bloom, N.; and Van Reenen, J. (2007); 'Constraints on Developing UK Management Practices.' Bloom, N.; Lemor, R.; Qi, M.; Sadun, R.; and Van Reenen, J. (2011).

133 'Leadership & management in the UK – The key to sustainable growth.' BIS (July 2012); page 44.

Table 4: Factors explaining historical variations in MFP growth rates across sectors, ranked by most important

Ranked … important factor by expert ➔ Factor ➔	most	second most	third most	fourth most	fifth most	Total mentions as one of the five most important factors	Total mentions as factor explaining UK MFP growth
Differences in scope for technological change	8	3	2	1	1	**15**	**16**
Differences in the extent of investment	5	2	0	4	0	**11**	**11**
Differences in scope for human capital gains	1	6	0	1	1	**9**	**9**
Interactions with place-based factors	3	2	3	0	0	**8**	**9**
Other factors	2	1	3	0	0	**6**	**6**
Differences in openness to international trade	1	2	1	1	0	**5**	**6**
Differences in appropriability of R&D investment	0	2	2	0	1	**5**	**5**
Differences in regulatory and competition policy	0	1	3	0	1	**5**	**5**
Barriers / frictions to labour substitutability	0	1	1	1	2	**5**	**7**
Differences in management performance	2	0	1	1	0	**4**	**5**
Differences in Government policy	0	0	2	1	1	**4**	**4**
Barriers / frictions to capital substitutability	1	1	1	0	0	**3**	**3**
Differences in firm ownership structure	1	0	0	2	0	**3**	**3**

Source: Economic Insight survey of academic experts, N=26.

The scope for technological change is intrinsic to industry characteristics

Of most relevance to the setting of OE at RIIO-3, it is critical to understand that the scope for technological change is a function of intrinsic industry characteristics. In other words, the main driver of across-sector variation in productivity (in the view of experts) is due to inherent differences in the features of said industries. The key explanations for this are as follows:

- Firstly, in some industries, the reliance on the latest technology / high tech as an input into their production processes is inherently higher than in others (i.e., more 'tech-heavy' industries). Therefore, compared to industries that intrinsically utilise technology assets to a lesser degree in the production process, improvements in technology *in general* will boost the productivity of inherently technological-focused industries more. For this reason, pharmaceuticals and computer manufacturing naturally have more scope to benefit from technological change than others.[134]

- Secondly, some industries' outputs themselves require a constant cycle of R&D investment and idea generation, as they require the constant generation of 'new products / services'. Pharmaceuticals would, again, be a good example of this, whereby new drugs and formulations are constantly required.

- Thirdly, asset lives vary materially across industries, which affects the speed / frequency with which they benefit from improvements in technology (taking the above two parameters as given). *Ceteris paribus*, firms that replace technology-related assets more frequently than others will benefit more from productivity related gains.

Unsurprisingly, the (minority of) industries that have performed more strongly in the UK in relation to productivity growth include said 'high-tech' industries. For example, between 1995 and 2019, average annual productivity growth for the following sectors was:

  - 12.95% for '*Telecommunications*';
  - 3.62% for '*Manufacture of computer, electronic and optical products*'; and
  - 1.16% for '*Manufacture of basic pharmaceutical products and pharmaceutical preparations*'.

In contrast to the above, the gas sector's intrinsic characteristics are <u>not</u> especially amenable to it having material scope to achieve large productivity gains from technology. Most evidently: (i) its main input factors are not 'high-tech' assets; (ii) it does not have to constantly develop and introduce new products or services in the way that pharmaceuticals has to; and (iii) it is characterised by long-lived assets.

[134] '*Factors behind cross-industry differences in technical progress*.' *Nelson, R. R. and Wolff, E. N. (1997).*

The scope for regulatory innovation incentives to materially impact productivity must be limited, as gains from technology are primarily related to industry characteristics

Regulators have suggested that innovation funding and incentives can improve productivity in regulated sectors. It has further been argued that this might enable regulated companies to *outperform* the current low levels of productivity growth in the UK. This proposition is misconceived, for three reasons.

- Firstly, where **the scope for industries to benefit (in productivity terms) from new technology is primarily a function of intrinsic industry characteristics,** then (by definition) the ability for any regulatory innovation incentive to materially affect productivity through this channel must be limited.

- Secondly, **the rationale for introducing a regulatory innovation incentive in the first place contradicts the conclusion that their existence provides a reason for regulated companies to outperform the UK economy on productivity.** Put simply, where there is a belief that innovation in said industries is 'inefficiently low' (i.e., because they are natural monopolies), the introduction of a regulatory innovation incentive, even if it were 100% effective, can only solve that market failure. That is, at best, it can bring innovation 'in line' with that which would occur without the market failure. Thus, innovation incentives provide reasons to consider regulated industries might perform *'more in line with'* the UK economy (relative to the counterfactual of them not existing). However, they do not provide a sound basis for believing regulated monopoly industries can materially outperform UK productivity growth.

'The materiality of innovation investment (in total, and under regulatory incentives) for gas networks is simply too low, relative to other industries, to have any meaningful effect on industry productivity.'

- Thirdly, **the materiality of innovation investment (in total, and under regulatory incentives) for gas networks is simply too low, relative to other industries,** to have any meaningful effect on industry productivity growth. This point is expanded below.

Ofgem provided gas networks with two sources of innovation funding under the RIIO-2 regulatory framework. We assess the materiality of each of these below.

- The strategic innovation fund is expected to deliver £450m of investment by 2026 across electricity and gas markets.[135] We do not know the exact amount going to gas networks, so have apportioned a share to gas distribution and transmission companies using:
  - the proportion of RIIO-2 totex allowances they account for (£12.7bn, 2018/19 prices), out of
  - the totex allowances for the electricity and gas transmission and distribution companies and the Electricity System Operator (£24.3bn).

Thus, we have assumed £234m (52% of £450m) can be expected to go to gas networks. However, this amounts to just 1.9% of gas networks' total allowances over RIIO-2.

[135] *'About the Strategic Innovation Fund (SIF).' Ofgem.*

- The network innovation allowance provides gas networks with £118m of network innovation funding over RIIO-2, comprising only 0.9% of their totex allowances.

Taken together, these innovation funding streams amount to just 2.8% of gas networks' totex allowances over RIIO-2. Other, more (intrinsically) innovative and high productivity industries, invest significantly more in R&D. Looking at these industries' proportion of total investment (GFCF) that is R&D over the last five years (2018-2020), we find that the proportion of total investment (GFCF) that is R&D amounts to:

  – 8.3% for '*Telecommunications*';

  – 65.5% for '*Manufacture of computer, electronic and optical products*'; and

  – 84.1% for '*Manufacture of basic pharmaceutical products and pharmaceutical preparations'.*

The large differences in the amounts invested in R&D across industries (with 'high-tech' industries having high proportions of investment in R&D) is further consistent with the intrinsic characteristics of industries determining their scope to benefit from technological change. In this context, regulatory innovation incentives and funding will likely not have any material impact on industry productivity. Note, this is not a criticism of those incentives, which may be effective, on their own terms.[136]

In summary: there is limited evidence that regulation can materially mitigate the impact of the slowdown on gas networks

Summarising the prior discussion, whilst we consider that regulation may (in principle) mitigate the impact of the slowdown on gas networks through 'certainty of investment', in practice the data is not supportive of this occurring. Furthermore, we find no reasons to believe regulation can materially impact gas networks' productivity (i.e., it cannot address wider issues with the human capital stock, management quality, and so on), once the wider set of (main) causal factors of the productivity growth flatline are considered. In addition, the evidence does not support the notion that innovation incentives and funding can materially mitigate the productivity growth slowdown in regulated industries.

[136] *A further consideration is the extent to which any regulatory innovation incentives tend to encourage productivity gains through 'cost reductions' or 'improved output / quality.' This issue matters in the historical context, whereby Ofgem has applied the entirety of the OE challenge to costs, whilst also requiring gas networks to meet certain output / outcome targets. We subsequently explain that such an approach results in a double count of OE and should be rectified. Should Ofgem retain that approach, however, then (even if innovation incentives did boost productivity) it would be necessary to identify and evidence 'to what degree' said incentives were relevant encouraging lower costs, in order for that to be the basis for adjusting an OE target applied entirely to company costs.*

4 Adjustments

As set out in Chapter 2, OE cannot be directly observed and it is therefore typically estimated using the benchmarking of productivity metrics (e.g. TFP). However, measures of productivity (such as TFP) will include productivity gains firms make from factors other than OE, such as catch-up efficiencies[137] and economies of scale. This arises from the fact that productivity metrics capture a *'change in output'* for a *'change in input',* but do not differentiate between the reasons for those changes. By basing OE estimates solely on unadjusted TFP, there is therefore the risk that the extent of OE would be incorrectly estimated (it could be over- or under-estimated).

Historically, there has therefore been considerable discussion as to whether *'raw'* estimates of OE (as implied by TFP benchmarking) may require post-benchmarking adjustments to correct for any disconnect between OE and TFP, which is not already addressed through other means. Accordingly, in this chapter we consider this issue.

Our view is that:

- **Effective comparator selection mitigates the risk of discrepancies between OE and TFP.** We mitigate a large part of the potential disconnect between TFP and OE through our data driven comparator selection methodology, as discussed in Chapter 2. The risk can be mitigated further by analysing the robustness of the resultant range to changes in analytical specifications, as we do in our sensitivity analyses in Annex 3.[138]

- **Post-benchmarking adjustments to OE should not be made**. The evidence, which we assess in this chapter, is ambiguous as to whether a net downward or upward adjustment should be made to our recommended OE range, based on potential differences between TFP and OE. This is because there are certain issues that may point to an upwards adjustment and others that point to a downwards one. However, in the main, the available evidence for each individual issue is only sufficient to (at most) establish its *directional* effect, but not to robustly *quantify its impact.* Therefore, one cannot determine whether the appropriate net impact of these factors would be that the benchmarked OE is under- / overstated. Thus, one cannot reliably determine whether a net upwards, or downwards, adjustment is appropriate.

137 *Ofgem set a catch-up efficiency challenge separately to OE in order to encourage the less efficient companies to 'catch-up' to the frontier company or companies as defined at the price control.*

138 *We do not recommend the use of these sensitivities to set the OE challenge for RIIO-3 because the sensitivities are based on assumptions that we do not recommend.*

- **The wider regulatory framework should be considered if any adjustments are made, to avoid double counting.** Previous discussion of post-benchmarking adjustments has also included the consideration of adjustments for overlaps with other elements of the regulatory framework (such as output incentives or measures of inflation). However, we consider that these overlaps should be resolved by ensuring the overall regulatory framework is consistently and robustly applied, rather than by adjusting the OE challenge itself.[139]

The remainder of this chapter addresses the following in turn:

- The evidence for adjustments arising from TFP capturing multiple efficiency savings, including: (i) catch-up efficiency; and (ii) economies of scale.
- The evidence for adjustments for embodied change.
- Overlaps with other elements of the regulatory framework (e.g. avoiding a double-count arising from outcomes / output incentives).

4A. TFP captures multiple efficiency savings

As discussed above, TFP may include efficiency gains over and above OE. In the following subsections, we therefore set out the evidence relating to this.

Catch-up efficiency

If an industry (or firm within an industry) is not perfectly efficient / competitive, by definition, its reported total productivity (e.g. TFP) will include both a degree of catch-up and OE related gains.

In choosing our set of comparators, we (partly) account for this, by using Criterion 2 to select industries that are more competitive, and therefore will have less gains from catch-up efficiency included in their TFP estimates. However, as no industry is perfectly competitive, this means that an OE challenge set using benchmarking data will still, strictly speaking, include some gains from catch-up efficiency, meaning that OE must be overstated. We note that removing the most highly concentrated industries from our preferred set (see Sensitivity 5[140]) causes our range to fall materially to from 0.2% to 0.8%, to -0.2% to 0.5%. This is consistent with the existence of upwards bias arising from catch-up efficiency.

Therefore, the inclusion of catch-up efficiency in TFP implies that it may be appropriate to make a downward adjustment to the OE range (or select a lower point in the range).

---

[139] *For example, if one's best view of OE for RIIO-3 is 0.5% pa and output / quality incentives are also set, then this 0.5% should be allocated between reduced cost and improved quality / output.*

[140] *We do not support using the results of this, or any of our other sensitivities, to inform the OE challenge. The sensitivities are based on assumptions that we do not recommend and are only included to test the robustness of our results. For example, Sensitivity 5 may inadvertently exclude industries that have a similar scope for economies of scale to gas networks by placing a greater weight on criterion 2 (the competitiveness of the industry).*

Economies of scale

Productivity metrics will also include gains industries and firms make from economies of scale. We account for this to some extent in our comparator selection, through our use of Criterion 3. However, no comparator set is going to perfectly replicate the scope for economies of scale for gas networks. Instead, the scope for gains from economies of scale will differ between the comparators and gas networks to some degree (leading to an over- *or* understatement of the OE challenge). Whether there is an over- or understatement due to economies of scale depends on whether the scope for gas networks to make gains from economies of scale is greater or less for the comparators used.

We note that the comparators in our preferred set likely have *slightly* less scope to make gains from economies of scale than the gas networks. It is not possible to measure this with precision but, to give an indication, we perform a high-level assessment. We do this by assessing how each comparator performs relative to gas networks on the metrics we use for Criterion 3. In below we show how each comparator compares to gas networks on Criteria 3a, 3b and 3c. For each of these, we mark the sector as "Red" if that measure implies it has *lower* scope for economies of scale than gas; "Amber" if that measure implies it has *similar* scope for economies of scale to gas;[141] and "Green" if that measure implies it has *greater* scope for economies of scale than gas. We then provide an *overall* assessment, as follows:

- The sector has *greater* scope than gas for gains from economies of scale. We define this as sectors that are ranked "Green" or "Amber" on all criteria;

- The sector has *less* scope than gas for gains from economies of scale. We define this sectors that are ranked "Red" or "Amber" on all criteria; or

- It is *unclear* whether the sector has more or less scope than gas for gains from economies of scale. We define this as sectors that are ranked "Green" on some criteria and "Red" on others.

We find that six of our comparator sectors are likely to have less scope than gas networks to make gains from economies of scale; and it is unclear for the remaining five sectors. This would imply that our results are *likely somewhat downward biased,* due to economies of scale, and that a *slight upward adjustment* to the OE range may be appropriate to account for this. However, we would caution drawing any firm conclusions from a high-level assessment that cannot quantify the impact of scale economies on productivity.

141 *We define this as where the divergence between the comparator and gas networks is less than 10%.*

Table 5: Economies of scale of comparators compared to gas networks

	3a: Tangible assets-turnover (difference from gas)	3b: Growth rate of tangible assets (difference from gas)	3c: Growth rate of turnover (difference from gas)	Overall assessment *(scope to make scale economies)*
Total industries (A-S)				**Lower**
Manufacturing				**Lower**
Chemicals; basic pharmaceutical products				**Unclear**
Manufacture of rubber and plastic products and other non-metallic mineral products				**Lower**
Computer, electronic, optical products; electrical equipment				**Unclear**
Manufacture of machinery and equipment n.e.c.				**Unclear**
Manufacture of motor vehicles, trailers, semi-trailers and of other transport equipment				**Lower**
Manufacture of furniture; jewellery, musical instruments, toys; repair and installation of machinery and equipment				**Lower**
Construction				**Unclear**
Wholesale and retail trade; repair of motor vehicles and motorcycles				**Lower**
Transportation and storage				**Unclear**

Source: Economic Insight analysis of FAME data

4B. Embodied technical change

Accounting for embodied change is especially challenging (and more so than for economies of scale and catch-up efficiency). This is because[142]:

- The exact *proportion* of embodied change included in the underlying TFP data is unclear (unlike for economies of scale and catch-up efficiency), but the evidence suggests that it is included to *some degree*.
- The extent to which it is included or excluded likely varies by industry.

As discussed in Section 2C, we consider that the analytical challenges that arise from this mean that it is more appropriate to address this issue by: (i) adding additional criteria in our sensitivity analyses[143] (we address this in Annex 3); and (ii) considering the role that post-benchmarking adjustments to our OE range might have in accounting for the extent of embodied change captured in our TFP estimates (we address this below).

There are four possible scenarios, which have different implications for adjustments to the range:

Table 6: Bias caused by embodied change in individual comparators

	Embodied technical change is fully (or mostly) captured by the comparator's TFP data	Embodied technical change is only partially captured by the comparator's TFP data
The comparator has similar scope to benefit from embodied change as gas networks	The comparator accurately represents the scope for gas networks to benefit from embodied change. No adjustment should be made.	The comparator's data downwards biases the OE estimate for gas networks and an upwards adjustment should be made.
The comparator has different scope to benefit from embodied change as gas networks	If the comparator has more scope to benefit from embodied change, the comparator's data upwards biases the OE estimate for gas networks and a downwards adjustment should be made. If the comparator has less scope to benefit from embodied change, the comparator's data downwards biases the OE estimate for gas networks and an upwards adjustment should be made.	If the comparator in fact, has more scope to benefit from embodied change than gas networks, but this is not fully captured in its data then it is possible that no adjustment should be made. However, it is possible that an upwards or downwards adjustment should be made, depending on the relative size of each effect (which cannot be reliably determined).

Source: Economic Insight

Note: We include an illustrative example in Annex 6

142 *We provide evidence for these facts in in Annex 6.*

143 *We do not recommend using the results of this sensitivity, or any other sensitivities used, to set the OE challenge for RIIO-3. This is because the sensitivities are based on assumptions that we do not recommend and are only included to test the robustness of our results. For example, including additional criteria related to embodied change may inadvertently exclude industries that are more similar to gas networks in other dimensions.*

In this case, we consider that:

- embodied technical change is more likely captured by our chosen comparators' TFP data than other comparator sets; and
- the comparators generally have more scope to benefit from embodied change than gas networks.

As shown in Table 6, this suggests that a downwards adjustment to the benchmarking OE range may be appropriate. However, we do not recommend that such an adjustment should be made, as the evidence is not sufficiently conclusive. Put simply, the relative size of the various effects cannot be determined, only the direction. This means we cannot confidently assign a robust 'size' to any adjustment. The relevant evidence is set out further below.

Embodied technical change is more likely captured by our chosen comparators' TFP data than other comparator sets

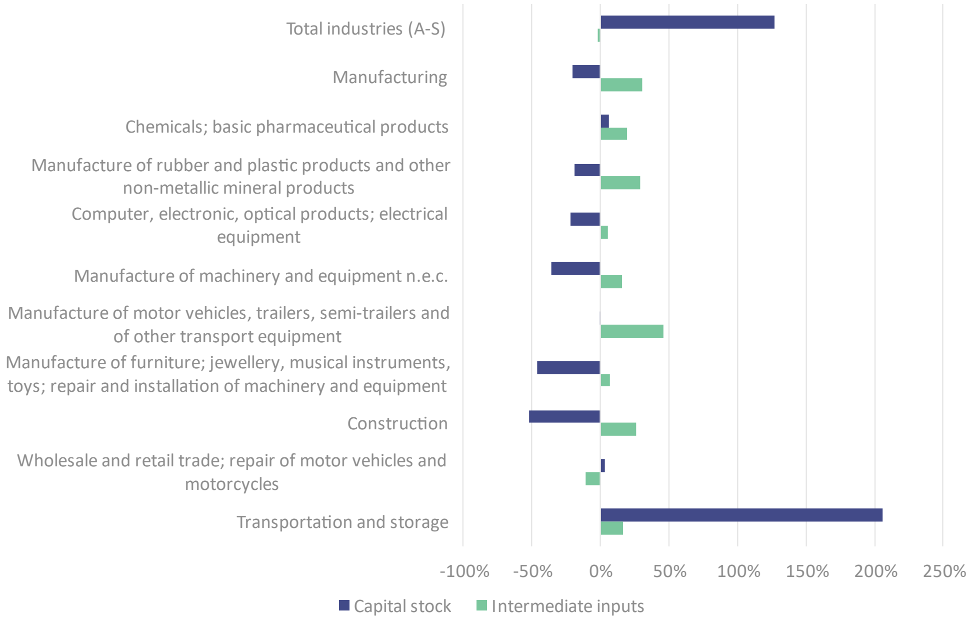
We can get an indication of the proportion of embodied technical change captured in the comparators' data by calculating the: (i) capital stock-GO ratio; and (ii) intermediate inputs-GO ratio for our comparator set.[144] This is because sectors that use *relatively*[145] more capital and intermediate inputs are likely to include a greater proportion of embodied change in their TFP data.[146] The following figure shows the result of this analysis for each industry in our preferred set.

[144] *Data on the intermediate inputs is from the tab "II_CP", available from the file "National Accounts" for the UK here: https://euklems-intanprod-llee.luiss.it/download/.*

[145] *We therefore inform our assessment by calculating the percentage difference between each comparator and the median industry for both metrics, using EU KLEMS data. The absolute amount of embodied change included in the data is not known.*

[146] *We explore the rationale behind this test further in Annex 6.*

Figure 16: Difference between average capital stock-gross output ratio, and intermediate inputs-gross output ratio, and the median across all industries (2013-2018)



Source: Economic Insight analysis of EU KLEMS data.

Taken together, the evidence in the figure above on intermediate inputs and capital intensity, suggests that the proportion of *included* (vs *excluded*) embodied change in the TFP of our comparators *may* be slightly greater than the UK industry average. We base this observation on the following:

- The data on capital intensity indicates that the proportion of embodied change included in the TFP data for our comparators is similar to the average industry. This is because there is a relatively even split between sectors in our preferred set that have capital stock-GO ratios above, or below, the median.[147]

- The data on intermediate inputs indicates that the proportion of embodied change included in the TFP data for our comparators is greater than the average industry. We observe that, on average, the comparators in our *'preferred set' generally* contain more intermediate inputs than the median industry.[148]

The comparators generally have more scope to benefit from embodied change than gas networks

As set out in our sensitivity analysis[149] (in Annex 3), industries that replace their assets at a similar rate will have a similar frequency of opportunities to benefit from any technology embedded in the new assets (all else equal). Gas networks are characterised

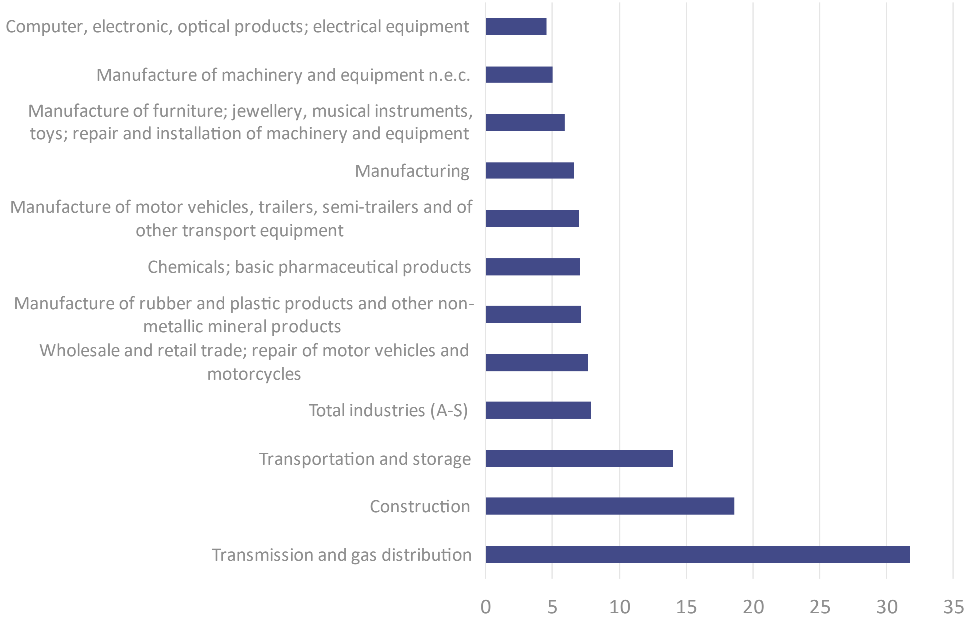
[147] *In particular, we note that: (i) two of our comparators have capital stock-GO ratios significantly greater than the median, (ii) three have ratios slightly greater than the median; (iii) three have ratios slightly less than the median; and (iv) three have ratios significantly less than the median (but the divergence is smaller in magnitude than the two sectors that have ratios significantly greater than the median).*

[148] *Nine of our eleven comparators contain greater levels of intermediate inputs than the median, and only two contain less.*

[149] *We do not recommended using the results of any of sensitivities to inform the OE challenge for RIIO-3. This is because the sensitivities are based on different assumptions and/or time periods to those we recommend.*

by long-lived assets, which are not replaced frequently. There is therefore more limited opportunity for them to benefit from new equipment with new technology (than industries with a higher asset replacement rate / utilising more high-tech assets). Indeed, the figure below demonstrates that the regulated networks have much longer asset lives than any of the comparators in our preferred set. Therefore, it is likely that the comparators in our preferred set have a greater scope to make gains from embodied technical change than the regulated networks themselves.

Figure 17: Average asset lives (years)



Source: Economic Insight analysis of company accounts data, aggregated through FAME.

Indeed, when we restrict our comparator set to those industries that: (i) have a relatively high proportion of embodied technical change included in their TFP estimates; and (ii) have a similar scope to gas networks to benefit from embodied change (see Sensitivity 1 in Annex 3 for more details), we find that our range shifts downward from 0.2%–0.8% to 0.1%–0.7%. The results of this sensitivity check are consistent with the view that the scope for gas networks to make gains from embodied technical change is likely low, in general, and relative to the sectors used in our analysis. However, we do not consider it appropriate to use this sensitivity (or any of our other sensitivities) to inform the OE challenge for RIIO-3. This is because the sensitivities are based on assumptions that we do not recommend and are only included to test the robustness of our recommended results. For example, we consider there to be too much uncertainty around the inclusion of embodied change in raw TFP data to account for it in our main results.

4C. Overlaps with other parts of the regulatory framework

OVERLAP ISSUES ARE COMPLEX AND GIVE RISE TO A RISK OF DOUBLE-COUNTING THE OE CHALLENGE.

Overlaps with other parts of the regulatory framework may point towards making a downward adjustment to the range, in order to avoid double-counting productivity gains because:

- Outcomes incentives set by Ofgem at RIIO-3 will account for some OE.
- Measures of inflation (e.g. CPIH, as used to index the RAV at RIIO-3) will reflect productivity gains.

However, we think the best way of dealing with the above is ensuring that the regulatory method more broadly is robust and consistently applied, once an appropriate overall OE challenge for gas networks has been identified (rather than being relevant to determining post-benchmarking adjustments and the appropriate OE challenge in the first place).

Output and outcomes incentives

The totality of any efficiency gains a firm (including gas networks) can make (which would include both catch-up and OE) can be realised through any combination of: (i) cost reductions; and (ii) increases in output (quality).

It therefore follows that, suppose (having undertaken the relevant analysis), one concluded that the appropriate OE challenge for gas networks was 1.0% pa. Assuming that this estimate was primarily derived using a comparator method, this means that this is the total productivity gain those comparator firms achieved, through any combination of cost savings / output (quality) improvements. Therefore, assuming the 1.0% is appropriate, to avoid any double-count risk, it would be important that the regulatory method at RIIO-3:

- allocates that 1.0% between cost reductions and output / outcome incentives; and
- avoids allocating the entirety of the OE challenge to cost reductions, *whilst also setting outputs / outcomes incentives that implicitly include a degree of OE challenge.*

In practice, as Ofgem is likely to set gas networks outputs / outcomes targets, the appropriate in principle approach is as follows:

- First, determine the total efficiency challenge *implicit* in any output / outcomes incentives that are set at RIIO-3 (i.e. improvements gas networks are asked to make out of base costs).
- Next, determine what proportion of the above is OE, as opposed to catch-up.
- Finally, only apply the *residual* amount of OE (i.e. that not captured in output / outcomes incentives) to company costs (i.e. as an OE cost reduction target).

By way of example, suppose that Ofgem's various outputs / outcomes incentives for gas networks were equivalent to an *implicit* OE challenge of 0.5% pa. That would leave a remaining 0.5% pa that could be applied to allowed costs as an OE cost reduction challenge, without a double-count occurring (assuming a total challenge of 1.0% pa was appropriate). Or, equivalently, if the OE cost challenge is set at 1.0% pa, but Ofgem also applied an implicit challenge via output and outcomes incentives, then the *implied overall* OE challenge (which accounts for both cost reductions and quality improvements) would be materially higher (at 1.5% pa).

Included in Annex 10 is a note from Professor Anthony Glass, confirming the above points. His key observations are as follows: *"[as] quality flows through to prices, we can interpret a value added-based output gain that is due to a revenue increase as reflecting an increase in the quantity supplied and / or an increase in the quality of the output… [and] an increase in gross output… of a comparator industry will typically reflect some combination of a reduction in production costs and an increase in quantity and / or quality. Relating this back to the setting of the annual frontier improvement targets for companies by domestic public utility regulators, the implication is that if the entirety of the annual frontier improvement targets are allocated to cost savings, but companies are simultaneously required to make quality improvements and /or possibly increase the quantity of at least one output, there will be a 'double-count'."*

Measures of inflation

Measures of inflation (CPIH, as used to index the RAV at RIIO-3) will reflect productivity gains. This follows from the fact that, as a broad measure of the price (inflation) consumers face, CPIH's main drivers are:

- changes in underlying costs incurred by companies when producing goods and services;
- changes in demand and supply; and
- efficiency savings (ergo productivity) firms make when producing goods and services.

There is therefore a potential for a tension in any regulatory method at RIIO-3 between Ofgem's approach to inflation (and real price effects) and its approach to productivity. Specifically:

- In relation to Ofgem's approach to inflation, this might be (broadly) characterised as being: compensation for general inflation risk (CPIH) is sufficient, save for some specifically identified instances, whereby allowances for real price effects will be made.
- However, under the above logic, gas networks should only be set an OE challenge equivalent to the difference between Ofgem's assessment of the industry specific OE and the OE *already captured in CPIH.*

The extent to which the above issue arises clearly depends on Ofgem's final approach to RPEs and inflation more broadly.

5 Conclusions and recommendations

In this final chapter, we summarise our: (i) key results; and (ii) wider recommendations as to the appropriate approach to OE at RIIO-3.

5A. Summary of our recommended OE range

Our estimates of the appropriate OE challenge at RIIO-3 are summarised in the table below. We consider that numbers towards the middle of the range are most appropriate.

Table 7: Summary of estimates for the OE challenge

	Recommended range	
	Low	High
OE estimate	0.2%	0.8%
Time period	2010-2019	Weighted average of: 1970-2007; and 1995-2019
Comparators	Preferred set	Preferred set

Source: Economic Insight analysis of EU KLEMS data

5B. Summary of our recommended approach

Our wider recommendations for how the estimation of OE should be approached at RIIO-3 are as follows.

The benchmarking approach should be transparent and robust

To avoid perceptions of *'cherry picking'*, any benchmarking approach at RIIO-3 should be transparent and robustly applied. Decisions on the key analytical choices must be evidence-based and allow for a consistent approach to be taken across price controls (to ensure that changes in OE over time likely reflect changes in underlying productivity potential, rather than being due to variations in method choices). Our views on how each of the key analytical decisions should be approached are as follows:

- **The productivity measure should be a total factor (TFP), rather than a partial factor (e.g. labour) metric, estimated on a GO (rather than VA) basis.** This is because the inclusion of partial factor metrics (such as labour productivity) wrongly assumes the TFP derived from benchmarking can be 'boosted' through a change in the mix of inputs (and in addition, TFP is more reflective of gas network company costs). Separately, GO is preferable to VA as a measure because:
  - it more accurately measures changes in productivity over time and across industries than VA (due to the inclusion of intermediate inputs); and

- it is less volatile and less subject to bias than VA.

- **Time period** choice should primarily be driven by ensuring internal consistency with other areas of the price control. Other important considerations are: (i) utilisation of the full data available; (ii) use of full business cycles; and (iii) the structural break in UK productivity growth.

- **Comparators** should be selected that: (i) reflect the scope for gas networks to make productivity gains; and (ii) ensure that the TFP estimates from these comparators are as applicable to an OE challenge as possible. To account for this, we recommend using the criteria outlined in this report. Furthermore, estimates of OE should be based on relatively wide comparator sets to: (i) avoid high variances in the annual TFP growth rates of individual sectors biasing the results; and (ii) reduce the sensitivity of the results to any updates to the data (rather than actual changes in productivity potential). Consistent with this, we think that some weight should always be placed on '*Total industries*', which exhibits much lower variance in its annual TFP growth rates than individual sectors.

The relevance of UK productivity growth to gas networks must be considered

We find limited reasons to suppose gas networks are materially shielded from the causal factors of the UK productivity growth slowdown since 2008:

- **The main factors causing the UK productivity growth slowdown are largely economy-wide and are unlikely to fully unwind over RIIO-3.** Evidence shows the key causal factors of the slowdown are insufficiency of: (i) investment; (ii) infrastructure quality; (iii) quality of human capital; and (iv) management quality.
- **Regulation is unlikely to mitigate the impact on gas networks of the factors causing the slowdown.** Regulation can only credibly mitigate the problem of underinvestment. However, data shows that, even on this issue, regulated industries are not insulated. As there are no strong reasons to believe regulation mitigates any of the other factors, its overall mitigating impact must be negligible.

Post-benchmarking adjustments should be avoided

There are various reasons why one might consider making post-estimation adjustments to a benchmarked OE estimate (e.g. TFP includes catch-up- and economies of scale-related gains; the question of to what degree embodied change is captured; etc.). Our report considers such issues in detail. However, we find that *'at best'* the directional impact of each issue can be identified and there is no reliable way to quantify their impact on OE. Therefore, one cannot determine whether the appropriate net impact of these factors would be that the benchmarked OE is under- / overstated. Thus, one cannot reliably determine whether a net upwards or downwards adjustment is appropriate. We thus do not recommend post-benchmarking adjustments to OE.

Point estimates from any benchmarked range should generally be taken from values 'towards the middle' of that range

As a point of principle, we consider best practice should be to derive any OE point estimate from towards the middle of any range derived from benchmarking. This reflects the inherent uncertainty as to the *'true'* value of OE, where it cannot be observed. It would be appropriate to depart from this if there were compelling evidence to the contrary (on a case-by-case basis). Again, this principle should help drive consistency over time and avoid accusations of cherry picking in either direction.

Care should be taken to avoid a 'double-count' of OE across other component of the price control

Finally, whatever the determined number for OE, it is critical to note that (as explained by Professor Anthony Glass – see Annex 10) this will be *inclusive* of productivity gains realised through improvements in quality / output. Therefore, when making its determinations, Ofgem should seek to 'allocate' the overall OE target between reduced costs and improved quality / output (rather than apply OE entirely to costs, whilst also setting outputs / outcomes incentives).

6 Annex 1 – Detailed results

In this annex, we provide further detail on the results of our *'recommended range'* that we present in Section 2C. In particular we present: (i) a detailed breakdown of our results showing the average productivity growth for each of our preferred sectors and time periods; and (ii) the results for our preferred set when alternative measures of productivity are used (we consider VA TFP, MFP and GO LEMS at constant capital). We only provide the results of alternative measures to GO TFP for completeness, since our view is that GO TFP is the most suitable measure for setting an OE challenge for gas networks (as we explain in Section 2A).

6A. Detailed breakdown of our recommended range results

The table below presents a detailed breakdown of our results for each sector in our preferred set, as well as each of our two recommended time periods.

Table 8: GO TFP detailed results for our preferred set

Comparator	2010-2019	Weighted average of: (i) 1970-2007; and (ii) 1995-2019
Final results (average)	0.2%	0.8%
Manufacturing	0.4%	0.9%
Chemicals; basic pharmaceutical products	1.2%	1.6%
Manufacture of rubber and plastic products and other non-metallic mineral products	1.0%	0.9%
Computer, electronic, optical products; electrical equipment	1.3%	2.2%
Manufacture of machinery and equipment n.e.c.	-0.2%	0.9%
Manufacture of motor vehicles, trailers, semi-trailers and of other transport equipment	-0.1%	0.7%
Manufacture of furniture; jewellery, musical instruments, toys; repair and installation of machinery and equipment	-0.4%	1.0%
Construction	-0.1%	0.1%
Wholesale and retail trade; repair of motor vehicles and motorcycles	-0.1%	-0.1%
Transportation and storage	-0.6%	0.5%
Total industries (A-S)	0.2%	0.2%

Source: Economic Insight analysis of EU KLEMS data

6B. Alternative measures of productivity

Below we present the results for our preferred comparator set if alternative measures of productivity were used. We consider these results inappropriate for informing the OE challenge at RIIO-3. This is because, as we set out in Section 2A, OE estimates should be based on TFP and GO. We only include the additional results below for completeness, not as alternatives to our recommended range.

VA TFP

The table below presents the results for our preferred set if TFP is measured in terms of VA rather than GO. These results are not appropriate for informing the OE challenge because VA is not an appropriate measure of OE as we detail in Section 2A.

Table 9: VA TFP results for preferred set

Comparator	2010-2019	Weighted average of: (i) 1970-2007; and (ii) 1995-2019
Final results (average)	0.5%	2.0%
Manufacturing	1.2%	2.4%
Chemicals; basic pharmaceutical products	2.7%	4.2%
Manufacture of rubber and plastic products and other non-metallic mineral products	2.8%	2.2%
Computer, electronic, optical products; electrical equipment	2.7%	5.6%
Manufacture of machinery and equipment n.e.c.	-0.6%	2.0%
Manufacture of motor vehicles, trailers, semi-trailers and of other transport equipment	-0.6%	2.3%
Manufacture of furniture; jewellery, musical instruments, toys; repair and installation of machinery and equipment	-0.9%	2.1%
Construction	-0.3%	0.2%
Wholesale and retail trade; repair of motor vehicles and motorcycles	-0.2%	0.0%
Transportation and storage	-1.5%	1.0%
Total industries (A-S)	0.3%	0.4%

Source: Economic Insight analysis of EU KLEMS data

ONS MFP

Another alternative measure of productivity growth is multifactor productivity (MFP), which we source from the ONS rather than EU KLEMS. MFP is defined by the ONS as the change in economic output that is not explained by changes in the inputs of quality adjusted labour and capital. The change in economic output is measured in terms of gross value added (GVA).[150] We present the results for our preferred set if MFP is used instead of GO TFP in the table below. We reiterate that we do not consider these results to be appropriate for informing the OE challenge. This is because MFP is a VA measure which is less appropriate than GO for gas networks, as we explain in Section 2A.

We also note that the ONS uses a different classification system for sectors than EU KLEMS. We therefore match the sectors in our preferred set from the EU KLEMS database to the closest equivalent sector in the ONS database. We show this matching in the table below. Where there is no equivalent sector in the ONS database, we display it as "N/A" in the table below and omit it from the calculation of the final results.

[150] *'A simple guide to multi-factor productivity.' ONS (2018)*

Table 10: ONS MFP results for our preferred set

EU KLEMS comparator	ONS equivalent comparator	2010-2019	1970-2019
Final results (average)		0.0%	1.0%
Manufacturing	Manufacturing	1.3%	2.5%
Chemicals; basic pharmaceutical products	N/A	N/A	N/A
Manufacture of rubber and plastic products and other non-metallic mineral products	N/A	N/A	N/A
Computer, electronic, optical products; electrical equipment	N/A	N/A	N/A
Manufacture of machinery and equipment n.e.c.	N/A	N/A	N/A
Manufacture of motor vehicles, trailers, semi-trailers and of other transport equipment	N/A	N/A	N/A
Manufacture of furniture; jewellery, musical instruments, toys; repair and installation of machinery and equipment	N/A	N/A	N/A
Construction	Construction	-0.5%	-0.5%
Wholesale and retail trade; repair of motor vehicles and motorcycles	Wholesale and retail trade; repair of motor vehicles and motorcycles	-0.1%	0.3%
Transportation and storage	Transportation and storage	-1.0%	1.9%
Total industries (A-S)	Total market sector	0.4%	1.0%

Source: Economic Insight analysis of ONS data

GO LEMS at constant capital

Instead of TFP, productivity growth could be measured using partial factor productivity measures. GO LEMS at constant capital is one such measure, and was used by CEPA at RIIO-2[151] and Ofgem at RIIO-1.[152] This measure differs from TFP as it assumes that capital is constant, and therefore only accounts for changes in the other inputs (i.e. labour and intermediate inputs). The measure is designed to eliminate the impact of capital substitution (i.e. a reduction in labour or intermediate inputs that has only been achieved by increasing the use of capital). We calculate it from EU KLEMS data using the following formula:[153]

> *GO LEMS at constant capital = GO TFP growth / share of labour and intermediate inputs in GO*[154]

The table below shows the results for our preferred set if this measure is used instead of GO TFP. Again, we do not consider these results to be appropriate for setting the OE challenge. This is because TFP is more appropriate than labour (or partial factor) productivity measures in this context as we set out in Section 2A.

---

[151] *'RIIO-GD2 and T2: Cost Assessment –Advice on Frontier Shift policy for Final Determinations.' CEPA (November 2020) table 1*

[152] *'RIIO-T1/GD1: Initial Proposals – Real price effects and ongoing efficiency appendix.' Ofgem (July 2012) page 18*

[153] *'RIIO-T1/GD1: Initial Proposals – Real price effects and ongoing efficiency appendix.' Ofgem (July 2012) page 18*

[154] *Share of labour and intermediate inputs in GO = (labour compensation + intermediate inputs) / GO*

Table 11: GO LEMS at constant capital results for our preferred set

Comparator	2010-2019	Weighted average of: (i) 1970-2007; and (ii) 1995-2019
Final results (average)	0.3%	0.9%
Manufacturing	0.5%	1.0%
Chemicals; basic pharmaceutical products	1.5%	1.9%
Manufacture of rubber and plastic products and other non-metallic mineral products	1.1%	1.0%
Computer, electronic, optical products; electrical equipment	1.5%	2.5%
Manufacture of machinery and equipment n.e.c.	-0.2%	1.0%
Manufacture of motor vehicles, trailers, semi-trailers and of other transport equipment	-0.1%	0.8%
Manufacture of furniture; jewellery, musical instruments, toys; repair and installation of machinery and equipment	-0.5%	1.1%
Construction	-0.1%	0.1%
Wholesale and retail trade; repair of motor vehicles and motorcycles	-0.1%	-0.1%
Transportation and storage	-0.7%	0.5%
Total industries (A-S)	0.2%	0.2%

Source: Economic Insight analysis of EU KLEMS data

7 Annex 2 – Comparator assessment

Table 12: Assessment of industries against criteria

Industry	Criteria								Preferred set	Sensitivity						In CEPA's narrow-set	In CEPA's economy-wide set
	1	2	3a	3b	3c	4a	4b	5		1	2	3	4	5	6		
Total industries									✓	✓	✗	✗	✗	✓	✓	✗	✗
Agriculture, forestry and fishing									✗	✗	✗	✗	✗	✗	✓	✗	✓
Mining and quarrying									✗	✗	✗	✗	✗	✗	✓	✗	✓
Manufacturing									✓	✓	✓	✓	✗	✗	✓	✗	✓
Manufacture of coke and refined petroleum products									✗	✗	✗	✗	✗	✗	✗	✗	✗
Chemicals; basic pharmaceutical products									✓	✗	✓	✓	✓	✗	✓	✗	✗
Manufacture of rubber and plastic products and other non-metallic mineral products									✓	✓	✓	✓	✓	✗	✓	✗	✗
Computer, electronic, optical products; electrical equipment									✓	✗	✓	✗	✓	✗	✓	✗	✗
Manufacture of machinery and equipment n.e.c.									✓	✗	✓	✓	✓	✗	✓	✗	✗
Manufacture of motor vehicles, trailers, semi-trailers and of other transport equipment									✓	✓	✓	✗	✓	✗	✓	✗	✗
Manufacture of furniture; jewellery, musical instruments, toys; repair and installation of machinery and equipment									✓	✗	✓	✗	✓	✗	✓	✗	✗

Industry	Criteria								Preferred set	Sensitivity						In CEPA's narrow-set	In CEPA's economy-wide set
	1	2	3a	3b	3c	4a	4b	5		1	2	3	4	5	6		
Electricity, gas, steam and air conditioning supply									✗	✗	✗	✗	✗	✗	✗	✗	✓
Water supply; sewerage, waste management and remediation activities									✗	✗	✗	✗	✗	✗	✗	✗	✓
Construction									✓	✓	✓	✗	✓	✓	✓	✓	✓
Wholesale and retail trade; repair of motor vehicles and motorcycles									✓	✗	✓	✗	✓	✓	✓	✓	✓
Transportation and storage									✓	✓	✓	✓	✓	✓	✓	✓	✓
Accommodation and food service activities									✗	✗	✗	✗	✗	✗	✗	✗	✓
Information and communication									✗	✗	✗	✗	✗	✗	✗	✗	✓
Financial and insurance activities									✗	✗	✗	✗	✗	✗	✗	✓	✓
Professional, scientific and technical activities; administrative and support service activities									✗	✗	✗	✗	✗	✗	✗	✗	✓
Arts, entertainment and recreation									✗	✗	✗	✗	✗	✗	✗	✗	✓
Other service activities									✗	✗	✗	✗	✗	✗	✗	✗	✓

Source: Economic Insight analysis

8 Annex 3 – Sensitivity analysis

This annex details the method we used to derive the sensitivity analysis that we discuss in Chapter 2 and Chapter 4. We do not consider the results of our sensitivities appropriate for informing the OE challenge at RIIO-3 because they are based on assumptions that we do not recommend. We only include the sensitivities to test the robustness of our recommended results and not as alternative estimates to our recommended range.

As part of our sensitivity analysis we test the robustness of our results to changing the time period over which the averages are calculated. We do this by:

- Including an additional period (1992-2007) to those we recommend (as we discuss in Section 2B).
- Calculating the average productivity growth over 1995-2019 (the whole NACE 2 period) and 1970-2007 (the whole NACE 1 period) separately, rather than combining the two in a weighted average. Calculating these separately will allow our sensitivity analysis to include more 'extreme points' than our recommended range. This is because the higher of 1995-2019 and 1970-2007 will be greater than the weighted average (assuming both are positive).

In addition to the time periods above, we continue to include the most recent business cycle (2010-2019) consistent with our recommended range. Therefore each sensitivity analysis will be based on the highest and lowest average across four time periods: (i) 1995-2019; (ii) 2010-2019; (iii) 1970-2007; and (iv) 1992-2007.

As well as changing the time period, we explore six further sets of comparators in addition to our preferred set. This allows us to test the robustness of our results to some of the assumptions we have made (again, recognising the subjectivity in comparator selection). We detail each of these sensitivities below, before providing our results for each sensitivity analysis.

8A. Selection of comparators for each sensitivity

Sensitivity 1 introduces two additional criteria to control for embodied technical change, while the other sensitivities are less complex and involve either: (i) strengthening or weakening our selection criteria; or (ii) adding or removing industries that are near the boundary for inclusion / exclusion under our criteria. The sensitivities are as follows:

Sensitivity 1: Embodied technical change

As we discuss in Section 2C, when measuring OE it is important that both embodied and disembodied technological change are included, in order for the full scope for productivity gains to be captured. Due to the uncertainty around the extent to which embodied technical change is included in raw TFP data, we do not account for it in the criteria when selecting our preferred set. Instead, we include a sensitivity to test the robustness of our results to accounting for embodied change (we also discuss the possibility of making a post-benchmarking adjustment for embodied technical change in Section 4B).

We consider that, if embodied technical change is only partly captured in TFP, to maximise the applicability of TFP estimates to an OE challenge, comparators should be selected that have:

(i) **A relatively higher proportion of productivity gains from embodied change included in their TFP estimates.** This ensures that the OE challenge for gas networks does not exclude embodied technical change in the first place; and

(ii) **A similar scope to gas networks to make gains from embodied change.** This ensures that the comparator's TFP estimates are an appropriate benchmark for gas networks (i.e. mitigates possible over or under estimation of OE).

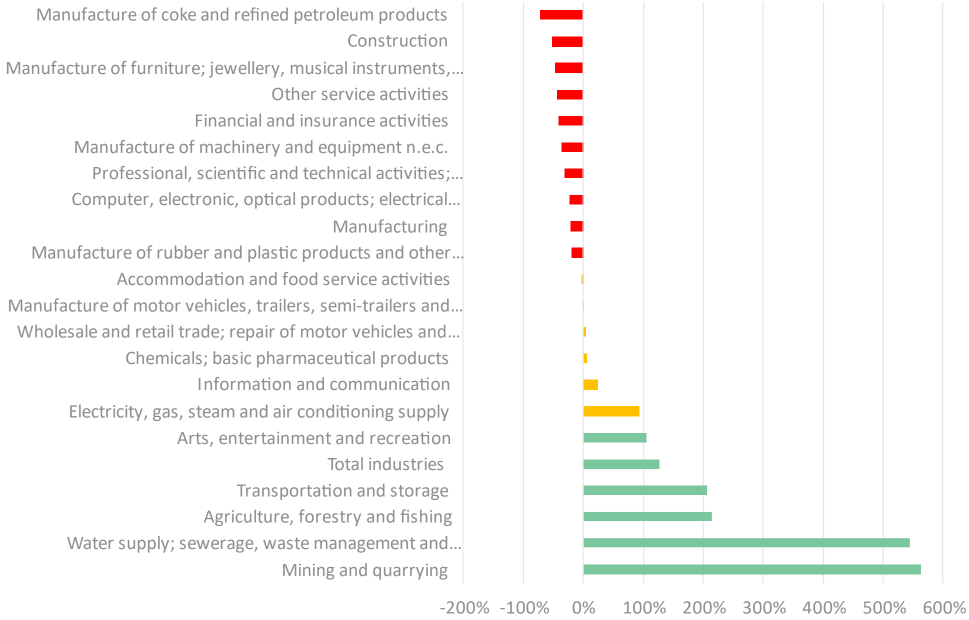
These considerations are explored in full detail in Table 6.

We therefore introduce additional sensitivity Criteria 4 and 5 to cover each of these points respectively – as follows:

(i) **Criteria 4.** Industries that: (i) are more capital intensive; and (ii) use *relatively* more intermediate inputs are likely to have a relatively higher proportion of gains from embodied technological change included within their TFP data (we explain why this is the case in Annex 6). We therefore include the following two criteria (calculated using the EU KLEMS data):

   – **Criteria 4a.** For this criteria, we calculate the average capital stock to gross output ratio from 2013-2018 for each potential comparator sector, and compare this to the median sector. We include this on the basis that comparator sectors with a higher ratio (relative to the median) are more capital intensive and therefore should have a relatively higher proportions of gains from embodied change included within their TFP data. We then assign the following three-colour scale: (i) "Green" where the absolute average divergence is greater than 100%; (ii) "Amber" where the absolute average divergence is between 0% and 100%; and (iii) "Red" where the absolute average divergence is less than 0%. The figure below shows the absolute divergence for each possible comparator sector.

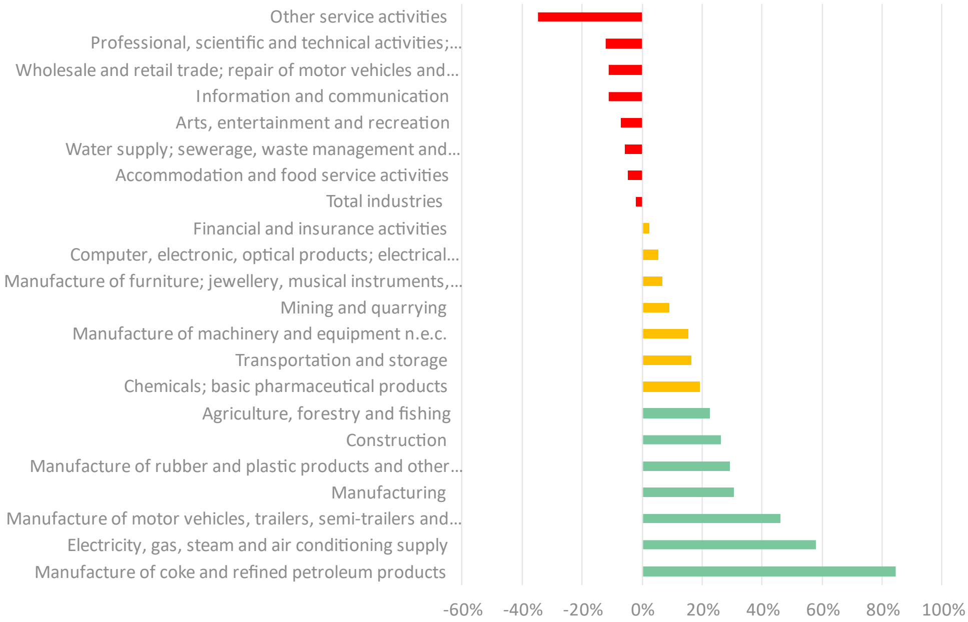
Figure 18: Difference between capital stock to gross output ratio, and the median across all industries (2013-2019)



Source: Economic Insight analysis of EU KLEMS data

– **Criteria 4b.** For this criteria, we calculate the average intermediate input to gross output ratio from 2013-2018 for each potential comparator, and compare this to the median sector. We include this on the basis that comparator sectors with a higher ratio should have a relatively higher proportion of gains from embodied technical change included within their TFP data. We then assign the three-colour scale, as follows: (i) "Green" where the absolute average divergence is greater than 20%; (ii) "Amber" where the absolute average divergence is between 0% and 20%; and (iii) "Red" where the absolute average divergence is less than 0%. The figure below shows the absolute divergence for each possible comparator sector.

Figure 19: Difference between average intermediate input to gross output ratio, and the median across all industries (2013-2018)

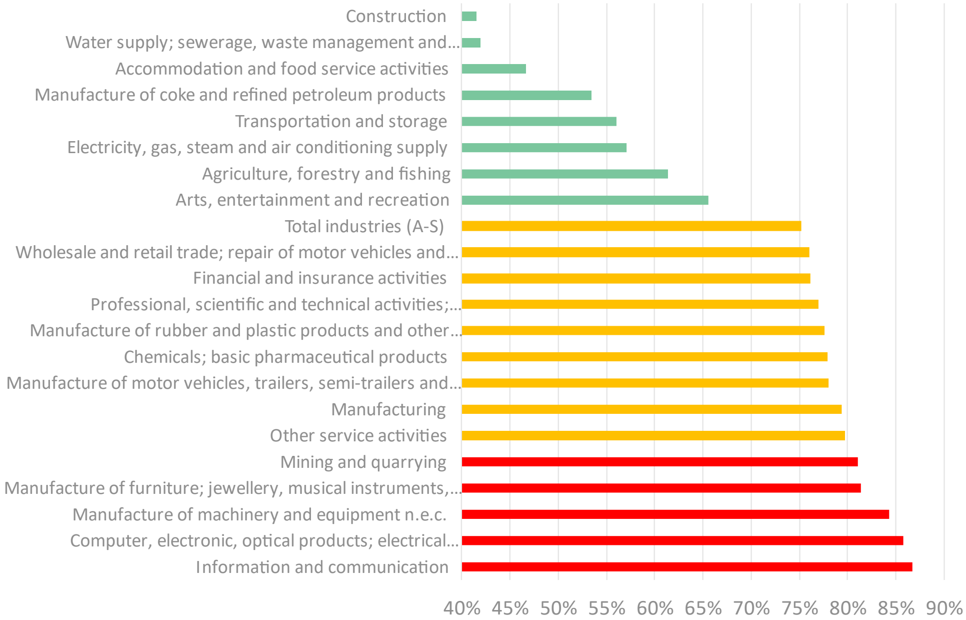


Source: Economic Insight analysis of EU KLEMS data

(ii) **Criteria 5.** There are limitations to the extent to which one can quantitatively evaluate the scope for making productivity gains from embodied change. However, for a given ability to benefit from new technology in production processes, the speed of asset replacement (asset life) will affect an industry's ability to utilise new technology. Therefore, we use similarity of asset lives as a metric (the rationale being that sectors with similar asset lives will be purchasing new equipment at a broadly similar rate). We calculate the absolute divergence in average asset life[155] between each sector and the gas networks. We then assign a three-colour scale, as follows: (i) "Green" where the absolute average divergence is less than 70%; (ii) "Amber" where the absolute average divergence is between 70% and 80%; and (iii) "Red" where the absolute average divergence is greater than 80%. The figure below shows the absolute divergence from the networks for each possible comparator sector.

155 *We calculate this from FAME data using the following formula: asset life = value of tangible fixed assets / annual depreciation. This gives the implied asset life in years, if we assume a linear rate of depreciation.*

Figure 20: Absolute divergence in asset lives (2013-2018)



Source: Economic insight analysis of FAME data

We then select the comparator set for this sensitivity using the following decision rule: (i) the comparators must be part of our preferred set; (ii) the comparators must be defined as "Green" in at least one of Criteria 4a and 4b to ensure their TFP measures contain relatively high levels of embodied change; and (iii) the comparators must also be defined as "Green" or "Amber" in Criterion 5, to ensure they have similar scope to make productivity gains from embodied change as gas networks.

Sensitivity 2: Remove total industries

This sensitivity includes all of the comparators in our preferred set, but we remove '*Total industries*'. This is to test the robustness of our results to its exclusion.

Sensitivity 3: Strengthen the threshold for similarity of economies of scale (and remove total industries)

In this sensitivity, as well as removing '*Total industries*', we strengthen Criterion 3. We do this to place a greater weight on the similarity of sectors to gas networks with regards to economies of scale. Specifically, this only includes comparators which: (i) meet all the criteria to be included in our preferred set; and (ii) at least two of Criteria 3a, 3b and 3c are ranked as "Green" or "Amber" (noting that, as for our preferred set, at least one of Criteria 3a, 3b and 3c must be ranked as "Green").

Sensitivity 4: Remove highly aggregated sectors

In this sensitivity, we remove all highly aggregated sectors (i.e. '*Total industries*' and '*Manufacturing*') from our preferred set, in order to test whether the implicit inclusion of sectors with some activities less similar to gas networks are affecting our results.

Sensitivity 5: Strengthen the threshold for competitive intensity

In this sensitivity, we strengthen Criterion 2 to test the effect of removing the least competitive comparators from our preferred set. We do this to further exclude the impact of catch-up efficiency on our results. Relative to our main results, we remove comparators that may be considered *'highly concentrated'* by the CMA[156] and only include comparators that are '*generally considered to be concentrated*'. In practice, this means that we remove all comparators from our preferred set for which the adjusted HHI in Figure 6 is greater than 2,000. We note that this has an equivalent effect on the chosen comparator set as removing all manufacturing sectors, which CEPA did at RIIO-2. CEPA omitted the manufacturing sectors because it believed that these sectors did not have similar enough activities to energy networks.

Sensitivity 6: Add '*Mining and quarrying*' and '*Agriculture, forestry and fishing*'

This sensitivity includes all of the comparators selected in our preferred set, plus '*Mining and quarrying*'; and '*Agriculture, forestry and fishing*'. This is because these sectors score relatively well on Criterion 2 and Criterion 3, but are ruled out by similarity of activities. We acknowledge that the similarity of activities criterion is inherently subjective to some extent and therefore we include these two industries to test the robustness of our results to weakening this criteria.

In Annex 2 we set out the comparators we include in each of the sensitivities (as well as our preferred set). We also show how each potential comparator performs on each of our criteria.

8B. Sensitivity analysis results

In this section we summarise the results from our sensitivity analysis, and then present a detailed breakdown of the results for each sensitivity in turn.

Summary

The results of our sensitivity analyses are summarised in the table below. This demonstrates what OE could be, under alternative comparator sets and time periods to those we recommend. We find that our results are robust to changes in our modelling approach (which we recognise is subjective to some extent), but emphasise that the approach used to generate our recommended range reflects the most balanced and well evidenced approach to estimating OE at RIIO-3. We therefore do not consider it appropriate to use the results of our sensitivities to inform the OE challenge at RIIO-3; we only include these sensitivities to test the robustness of our main results. Relatedly, we do not consider that any of these sensitivities reflect a "better" approach than other sensitivities, rather they simply test the robustness of our results to different tweaks to the specification. We base our sensitivities on: (a) our preferred comparator set and our sensitivity comparator sets; and (b) the time periods: 1995-2019, 2010-2019; 1970-2007; and 1992-2007.

[156] *The State of UK Competition,' CMA (April 2022); paragraph 2.10.*

Table 13: Summary of results from sensitivity analysis

Comparator set	Average GO TFP growth (%)	-0.2	-0.1	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Recommended range	0.2% to 0.8%													
Preferred set (with additional time periods)	0.2% to 0.9%													
Sensitivity 1: Embodied technical change	0.1% to 0.7%													
Sensitivity 2: Remove total industries	0.2% to 0.9%													
Sensitivity 3: Strengthen threshold for similarity of economies of scale	0.3% to 1.0%													
Sensitivity 4: Remove highly aggregated sectors	0.2% to 0.9%													
Sensitivity 5: Strengthen threshold for intensity of competition	-0.2% to 0.5%													
Sensitivity 6: Add '*Mining and quarrying*'; and '*Agriculture, forestry and fishing*'	0.1% to 0.7%													

Source: Economic Insight analysis of EU KLEMS data

Note: The midpoint of each range is highlighted in a different colour

Detailed sensitivity results

The tables below present a detailed breakdown of the results for each sensitivity analysis in turn. These provide detail on the average productivity growth for each sector and time period within each sensitivity.

Table 14: Results for preferred set with additional time periods

Comparator	1995-2019	2010-2019	1970-2007	1992-2007
Final results (average)	0.9%	0.2%	0.8%	0.7%
Manufacturing	1.1%	0.4%	0.7%	0.6%
Chemicals; basic pharmaceutical products	1.9%	1.2%	1.3%	1.1%
Manufacture of rubber and plastic products and other non-metallic mineral products	0.9%	1.0%	0.9%	0.4%
Computer, electronic, optical products; electrical equipment	2.9%	1.3%	1.7%	1.8%
Manufacture of machinery and equipment n.e.c.	1.3%	-0.2%	0.5%	1.0%
Manufacture of motor vehicles, trailers, semi-trailers and of other transport equipment	0.7%	-0.1%	0.8%	0.5%
Manufacture of furniture; jewellery, musical instruments, toys; repair and installation of machinery and equipment	1.0%	-0.4%		
Construction	-0.2%	-0.1%	0.3%	0.1%
Wholesale and retail trade; repair of motor vehicles and motorcycles	-0.2%	-0.1%	0.1%	0.7%
Transportation and storage	-0.1%	-0.6%	1.1%	0.9%
Total industries (A-S)	0.2%	0.2%	0.2%	0.3%

Source: Economic Insight analysis of EU KLEMS data

Table 15: Results for Sensitivity 1: Embodied technical change

Comparator	1995-2019	2010-2019	1970-2007	1992-2007
Final results (average)	0.4%	0.1%	0.7%	0.5%
Manufacturing	1.1%	0.4%	0.7%	0.6%
Manufacture of rubber and plastic products and other non-metallic mineral products	0.9%	1.0%	0.9%	0.4%
Manufacture of motor vehicles, trailers, semi-trailers and of other transport equipment	0.7%	-0.1%	0.8%	0.5%
Construction	-0.2%	-0.1%	0.3%	0.1%
Transportation and storage	-0.1%	-0.6%	1.1%	0.9%
Total industries (A-S)	0.2%	0.2%	0.2%	0.3%

Source: Economic Insight analysis of EU KLEMS data

Table 16: Results for Sensitivity 2: Remove total industries

Comparator	1995-2019	2010-2019	1970-2007	1992-2007
Final results (average)	0.9%	0.2%	0.8%	0.8%
Manufacturing	1.1%	0.4%	0.7%	0.6%
Chemicals; basic pharmaceutical products	1.9%	1.2%	1.3%	1.1%
Manufacture of rubber and plastic products and other non-metallic mineral products	0.9%	1.0%	0.9%	0.4%
Computer, electronic, optical products; electrical equipment	2.9%	1.3%	1.7%	1.8%
Manufacture of machinery and equipment n.e.c.	1.3%	-0.2%	0.5%	1.0%
Manufacture of motor vehicles, trailers, semi-trailers and of other transport equipment	0.7%	-0.1%	0.8%	0.5%
Manufacture of furniture; jewellery, musical instruments, toys; repair and installation of machinery and equipment	1.0%	-0.4%		
Construction	-0.2%	-0.1%	0.3%	0.1%
Wholesale and retail trade; repair of motor vehicles and motorcycles	-0.2%	-0.1%	0.1%	0.7%
Transportation and storage	-0.1%	-0.6%	1.1%	0.9%

Source: Economic Insight analysis of EU KLEMS data

Table 17: Results for Sensitivity 3: Strengthen threshold for similarity of economies of scale

Comparator	1995-2019	2010-2019	1970-2007	1992-2007
Final results (average)	1.0%	0.3%	0.9%	0.8%
Manufacturing	1.1%	0.4%	0.7%	0.6%
Chemicals; basic pharmaceutical products	1.9%	1.2%	1.3%	1.1%
Manufacture of rubber and plastic products and other non-metallic mineral products	0.9%	1.0%	0.9%	0.4%
Manufacture of machinery and equipment n.e.c.	1.3%	-0.2%	0.5%	1.0%
Transportation and storage	-0.1%	-0.6%	1.1%	0.9%

Source: Economic Insight analysis of EU KLEMS data

Table 18: Results for Sensitivity 4: Remove highly aggregated sectors

Comparator	1995-2019	2010-2019	1970-2007	1992-2007
Final results (average)	0.9%	0.2%	0.8%	0.8%
Chemicals; basic pharmaceutical products	1.9%	1.2%	1.3%	1.1%
Manufacture of rubber and plastic products and other non-metallic mineral products	0.9%	1.0%	0.9%	0.4%
Computer, electronic, optical products; electrical equipment	2.9%	1.3%	1.7%	1.8%
Manufacture of machinery and equipment n.e.c.	1.3%	-0.2%	0.5%	1.0%
Manufacture of motor vehicles, trailers, semi-trailers and of other transport equipment	0.7%	-0.1%	0.8%	0.5%
Manufacture of furniture; jewellery, musical instruments, toys; repair and installation of machinery and equipment	1.0%	-0.4%		
Construction	-0.2%	-0.1%	0.3%	0.1%
Wholesale and retail trade; repair of motor vehicles and motorcycles	-0.2%	-0.1%	0.1%	0.7%
Transportation and storage	-0.1%	-0.6%	1.1%	0.9%

Source: Economic Insight analysis of EU KLEMS data

Table 19: Results for Sensitivity 5: Strengthen threshold for intensity of competition

Comparator	1995-2019	2010-2019	1970-2007	1992-2007
Final results (average)	-0.1%	-0.2%	0.4%	0.5%
Construction	-0.2%	-0.1%	0.3%	0.1%
Wholesale and retail trade; repair of motor vehicles and motorcycles	-0.2%	-0.1%	0.1%	0.7%
Transportation and storage	-0.1%	-0.6%	1.1%	0.9%
Total industries (A-S)	0.2%	0.2%	0.2%	0.3%

Source: Economic Insight analysis of EU KLEMS data

Table 20: Results for Sensitivity 6: Add *'Mining and quarrying'*; and *'Agriculture, forestry and fishing'*

Comparator	1995-2019	2010-2019	1970-2007	1992-2007
Final results (average)	0.6%	0.1%	0.7%	0.7%
Agriculture, forestry and fishing	1.2%	1.1%	1.0%	0.6%
Mining and quarrying	-2.3%	-1.7%	-0.7%	0.3%
Manufacturing	1.1%	0.4%	0.7%	0.6%
Chemicals; basic pharmaceutical products	1.9%	1.2%	1.3%	1.1%
Manufacture of rubber and plastic products and other non-metallic mineral products	0.9%	1.0%	0.9%	0.4%
Computer, electronic, optical products; electrical equipment	2.9%	1.3%	1.7%	1.8%
Manufacture of machinery and equipment n.e.c.	1.3%	-0.2%	0.5%	1.0%
Manufacture of motor vehicles, trailers, semi-trailers and of other transport equipment	0.7%	-0.1%	0.8%	0.5%
Manufacture of furniture; jewellery, musical instruments, toys; repair and installation of machinery and equipment	1.0%	-0.4%		
Construction	-0.2%	-0.1%	0.3%	0.1%
Wholesale and retail trade; repair of motor vehicles and motorcycles	-0.2%	-0.1%	0.1%	0.7%
Transportation and storage	-0.1%	-0.6%	1.1%	0.9%
Total industries (A-S)	0.2%	0.2%	0.2%	0.3%

Source: Economic Insight analysis of EU KLEMS data

9 Annex 4 – Update to CEPA's RIIO-2 method

This annex details our update of CEPA's RIIO-2 OE benchmarking analysis to the latest EU KLEMS release. At RIIO-2 CEPA calculated its results using the EU KLEMS 2019 release but, since then, EU KLEMS have released an updated version of this dataset (the 2023 release). We summarise our results from updating CEPA's method to use this data in Section 2C. In this annex we provide further details by:

- **Outlining the method used by CEPA at RIIO-2** and any differences between this and our update of CEPA's results.

- **Replicating CEPA's RIIO-2 results** using the same data and method to CEPA. We demonstrate that our replication gives nearly identical results to CEPA, suggesting that our replication is accurate. The magnitude of any differences between our results and CEPA's are sufficiently small (less than 0.07 percentage points), which suggests that the discrepancies are most likely attributable to rounding errors at some point during the calculation process.

- **Performing a *'straight update'* of CEPA's RIIO-2 method to the latest release of EU KLEMS** (while keeping the rest of CEPA's method identical to that used at RIIO-2). We demonstrate that this *'straight update'* changes the results for CEPA's targeted set dramatically, while its results for the economy-wide set are less sensitive.

- **Performing a *'complete update'* of CEPA's RIIO-2 method to the latest time period and release of EU KLEMS** (while keeping the rest of CEPA's method identical to that used at RIIO-2). We demonstrate that updating the time period to include the full sample now available, only has a small impact on CEPA's results (relative to updating to the latest release of EU KLEMS). However, both updating the time period, and updating the range cause a reduction in the implied OE range estimated using CEPA's method.

We do not support using a 'straight update' or 'complete update' of CEPA's method to set the OE challenge for RIIO-3 because, as we set out in the Chapter 1, we believe that this approach was not robust.

9A. CEPA's method

CEPA followed a benchmarking approach using EU KLEMS data, similar to the one we use to derive our results in this report (we describe any differences between our approach and CEPA's in Annex 5). CEPA's method provided a range of eight historical productivity measures that it recommended Ofgem should take into account when setting the OE challenge.[157] We present the key aspects of CEPA's method and any differences to our replication and updates in the table below. Blank cells indicate we have taken the same approach as CEPA.

[157] *'RIIO-GD2 and T2: Cost Assessment –Advice on Frontier Shift policy for Final Determinations.' CEPA (November 2020) page 6*

Table 21: Key aspects of CEPA's method

	CEPA's RIIO-2 method	Our replication of CEPA's results	Our '*straight update*' of CEPA's results to the latest data	Our '*complete update*' of CEPA's results to the latest time period and data
Dataset	EU KLEMS 2019 release[158]		EU KLEMS 2023 release	EU KLEMS 2023 release
Time period	1997-2016[159]			1995-2019
Productivity measures	OE range based on four difference measures:[160] (1) VA TFP (2) GO TFP[161] (3) VA LP at constant capital[162] (4) GO LEMS at constant capital[163]			
Comparators	OE range based on two different comparator sets:[164] (1) Targeted (narrow) set[165] (2) Weighted average[166] of economy-wide set[167]			

Source: CEPA and Ofgem RIIO-2 documents detailed in the footnotes below

Notes: Blank cells indicate we have taken the same approach as CEPA when replicating or updating its results.

158 'RIIO-GD2 and T2: Cost Assessment –Advice on Frontier Shift policy for Final Determinations.' CEPA (November 2020) page 6

159 'RIIO-GD2 and T2: Cost Assessment –Advice on Frontier Shift policy for Final Determinations.' CEPA (November 2020) page 6

160 'RIIO-GD2 and T2: Cost Assessment –Advice on Frontier Shift policy for Final Determinations.' CEPA (November 2020) page 6

161 CEPA calculated this using the following formula: GO TFP growth = VA TFP growth x (VA / GO). See: 'RIIO-GD2 and T2: Cost Assessment – Frontier shift methodology paper.' CEPA (May 2020) page 14

162 We assume CEPA used the approach taken by Ofgem at RIIO-1, which calculated this measure from EU KLEMS data using the following formula: VA LP at constant capital = VA TFP growth / share of labour in VA. See: 'RIIO-T1/GD1: Initial Proposals – Real price effects and ongoing efficiency appendix.' Ofgem (July 2012) page 17

163 CEPA stated that this was a metric used by Ofgem at RIIO-1, which calculated this measure using the following formula: GO LEMS at constant capital = GO TFP growth / share of labour and intermediate inputs in GO. 'RIIO-T1/GD1: Initial Proposals – Real price effects and ongoing efficiency appendix.' Ofgem (July 2012) page 18

164 'RIIO-GD2 and T2: Cost Assessment –Advice on Frontier Shift policy for Final Determinations.' CEPA (November 2020) page 6

165 This was made up of the following four sectors: 'Construction', 'Wholesale and retail trade: repair of motor vehicles and motorcycles'; 'Transportation and storage'; and 'Financial and insurance activities'.

166 CEPA weighted each industry by the share of GO or VA (corresponding to the productivity measure used). Our replication of CEPA's methodology indicates that it based these shares on the most recent year in the time period used i.e. for the period 1997-2016, the share would be based on 2016.

167 This comparator set contains the following industries: 'Agriculture, forestry and fishing'; 'Mining and quarrying'; ' Total Manufacturing'; 'Electricity, gas, steam and air conditioning supply'; 'Water supply; sewerage, waste management and remediation activities'; 'Construction'; 'Wholesale and retail trade; repair of motor vehicles and motorcycles'; 'Transportation and storage'; 'Accommodation and food service activities'; 'Information and communication'; 'Financial and insurance activities'; 'Professional, scientific and technical activities; administrative and support service activities'; 'Arts, entertainment and recreation'; and 'Other service activities'. See: 'RIIO-GD2 and T2: Cost Assessment – Frontier shift methodology paper.' CEPA (May 2020) table A.5

Our replication of CEPA's methodology indicates that a geometric mean was used to average productivity growth across time (although this was not explicitly stated by CEPA). We therefore use a geometric mean in our replication and updates of CEPA's results. We detail the difference between arithmetic and geometric means in Annex 8.

9B. Replication of CEPA's RIIO-2 results

We replicate CEPA's results following the method it used at RIIO-2 (as detailed above). As shown in the table below, our results are nearly identical to those reported by CEPA in its final determination report for RIIO-2. In fact, six of the eight figures are identical to CEPA's reported results (which it reports to one decimal place). The remaining two figures only differ from CEPA's results by less than 0.07 percentage points. This suggests that our replication of CEPA's methodology is accurate, because any differences are so small that they are most likely the result of rounding errors during the calculation process.

Table 22: Replication of CEPA's RIIO-2 results

Measure	Expenditure category	Targeted comparator set		Economy-wide comparator set (weighted)	
		CEPA	EI replication	CEPA	EI replication
VA LP at constant K	opex	0.8%	**0.8%**	1.0%	**0.93%**[168]
VA TFP	capex, repex, opex	0.5%	**0.5%**	0.9%	**0.9%**
GO LEMS at constant K	opex	0.3%	**0.3%**	0.5%	**0.5%**
GO TFP	capex, repex, opex	0.2%	**0.26%**[169]	0.4%	**0.4%**

Source: CEPA's final determination report at RIIO-2[170] and Economic Insight analysis of EU KLEMS data

[168] *We report this to two decimal places because it is one of only two figures that is not quite equal to CEPA's reported results. Reporting to two decimal places makes possible to observe the magnitude of the difference between CEPA's reported results and our replication to a greater level of granularity.*

[169] *We report this to two decimal places because it is one of only two figures that is not quite equal to CEPA's reported results. Reporting to two decimal places makes it possible to observe the magnitude of the difference between CEPA's reported results and our replication to a greater level of granularity.*

[170] *'RIIO-GD2 and T2: Cost Assessment –Advice on Frontier Shift policy for Final Determinations.' CEPA (November 2020) table 2.1*

9C. Updating CEPA's results to the latest EU KLEMS release

The table below compares CEPA's reported RIIO-2 results to our *'straight update'* of its analysis which uses the latest EU KLEMS release (while keeping all other parts of the method identical to CEPA's RIIO-2 approach). As we note in Section 2C the OE range for CEPA's targeted comparator set undergoes a marked shift from 0.2% to 0.8% (with the old data), to -0.8% to -0.2% (with the latest data). In contrast, the OE efficiency range for CEPA's economy-wide comparator set is less sensitive. It changes from 0.4% to 1.0% with the old data, to 0.4% to 0.8% with the new data.

Table 23: *'Straight update'* of CEPA's RIIO-2 method to the latest EU KLEMS release

Measure	Expenditure category	Targeted comparator set		Economy-wide comparator set (weighted)	
		CEPA	EI update	CEPA	EI update
VA LP at constant K	opex	0.8%	**-0.8%**	1.0%	**0.5%**
VA TFP	capex, repex, opex	0.5%	**-0.6%**	0.9%	**0.8%**
GO LEMS at constant K	opex	0.3%	**-0.2%**	0.5%	**0.4%**
GO TFP	capex, repex, opex	0.2%	**-0.2%**	0.4%	**0.4%**

Source: CEPA's final determination report at RIIO-2[171] and Economic Insight analysis of EU KLEMS data

The results of updating the time period to the full sample is less dramatic, as we demonstrate in the table below. The table compares: (i) our *'straight update'* of CEPA's analysis using the period 1997-2016 and the latest EU KLEMS release; to (ii) our *'complete update'* of CEPA's analysis that changes the time period to 1995-2019 (as well as updating to the latest EU KLEMS release). We update the time period to allow use of the wider sample that is now available following the most recent EU KLEMS release. Using the full time period available is also consistent with CEPA's approach at RIIO-ED2 for the draft[172] and final determinations.[173]

Updating the time period from 1997-2016 to 1995-2019 only has a small effect on the implied OE range relative to updating the data. For the targeted comparator set, the productivity growth range changes from -0.2% to -0.8% (for the period 1997-2016) to -0.2% to -1.1% (for 1995-2019). For the economy-wide comparator set, the range changes from 0.4% to 0.8% (for 1997-2016), to 0.2% to 0.5% (for 1995-2019).

171 *'RIIO-GD2 and T2: Cost Assessment –Advice on Frontier Shift policy for Final Determinations.' CEPA (November 2020) table 2.1*
172 *'RIIO-ED2: Cost Assessment – Frontier Shift methodology paper.' CEPA (June 2022) page 36*
173 *'RIIO-ED2 Final Determinations Core Methodology Document.' Ofgem (November 2022) paragraph 7.630*

Table 24: *'Complete update'* of CEPA's RIIO-2 method to latest EU KLEMS release and time period

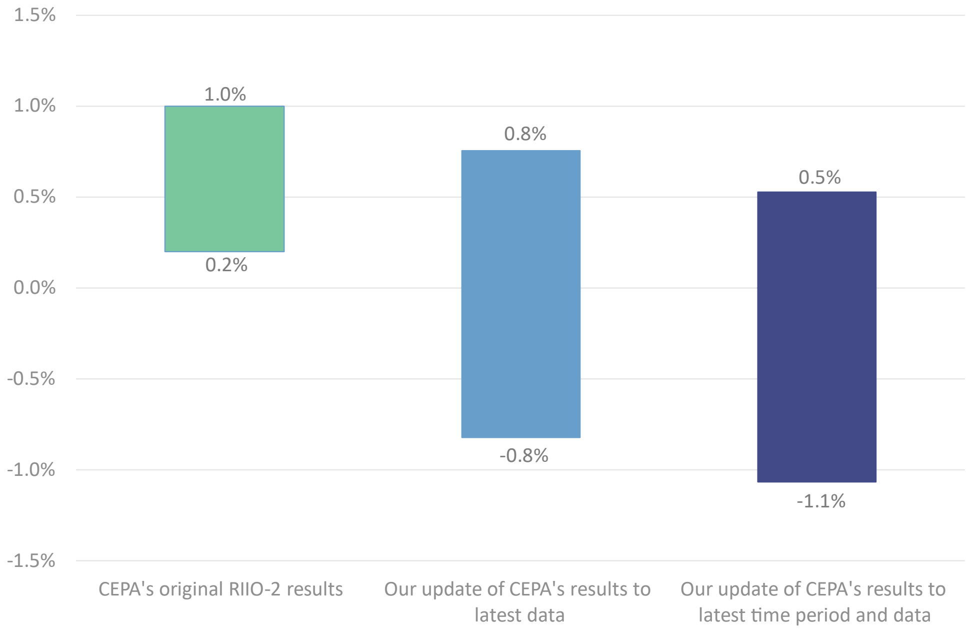
Measure	Expenditure category	Targeted comparator set		Economy-wide comparator set (weighted)	
		'straight update'	*'complete update'*	*'straight update'*	*'complete update'*
VA LP at constant K	opex	-0.8%	**-1.1%**	0.5%	**0.2%**
VA TFP	capex, repex, opex	-0.6%	**-0.5%**	0.8%	**0.5%**
GO LEMS at constant K	opex	-0.2%	**-0.3%**	0.4%	**0.3%**
GO TFP	capex, repex, opex	-0.2%	**-0.2%**	0.4%	**0.3%**

Source: Economic Insight analysis of EU KLEMS data

The figure below summarises the impact of each of the changes on the implied OE range. It shows the range for: (i) CEPA's original reported results; (ii) our update of the analysis to the latest data release; and (iii) our update of the analysis to change the time period as well as the underlying data. Both updating the dataset used and the time period covered cause a downward shift in the overall range given by CEPA's method.[174] However, updating the dataset has a more material effect than updating the time period.

[174] *The overall range is determined by taking the highest and lowest values of the eight point estimates given by CEPA's method.*

Figure 21: Summary of OE ranges based on CEPA's RIIO-2 methodology



Source: CEPA's final determination report at RIIO-2[175] and Economic Insight analysis of EU KLEMS data

[175] *'RIIO-GD2 and T2: Cost Assessment –Advice on Frontier Shift policy for Final Determinations.' CEPA (November 2020) table 2.1*

10 Annex 5 – Comparison to CEPA's RIIO-2 method

In this annex we compare the key analytical decisions made in our benchmarking analysis to those made by CEPA at RIIO-2. We note that our method is not intended to build upon or augment CEPA's approach, rather our approach is guided by the three principles set out in Chapter 1 and is consistent with best practice. As we set out in the Chapter 1, this is to avoid fundamental issues with a re-run of the RIIO-2 approach, which was not robust. Nonetheless, for completeness, Table 25 below summarises the key differences between our approach and CEPA's RIIO-2 approach, and our reasoning for our decisions. We then compare our comparator set to CEPA's two comparator sets in Table 26, and describe the differences between the sets.

Table 25: Comparison of our key analytical decisions to CEPA's RIIO-2 approach

Analytical decision	Our choice	CEPA's choice	Our reasoning
Dataset	EU KLEMS 2023 release	EU KLEMS 2019 release	There has been a updated dataset released since RIIO-2. We use the most up-to-date data available.
TFP vs LP	TFP	TFP only for capex & repex. Both TFP and LP for opex.	The proportion of non-labour costs in opex are too high for labour productivity to be reflective of the potential for productivity gains. We provide more detail for our choice in Section 2A.
GO vs VA	GO	Both	GO is more appropriate than VA because it: (i) is more accurate that VA; (ii) avoids the bias in VA created by exclusion of intermediate inputs (which are relevant for gas networks); and (iii) is preferred for industry specific productivity, as shown in the literature. We detail these arguments in Section 2A.
Time period	(1) 2010-2019 (2) Weighted average of: (i) 1995-2019; and (ii) 1970-2007	1997-2016	Our first period provides the likely lower bound given by a persistence of the slow productivity growth in the recent past. CEPA's range ignores this possibility. Our second period provides a likely upper bound by allowing for a *'partial unwinding'* of the productivity slowdown. It utilises the largest possible sample, which is larger than that used by CEPA. We detail our choice of time period in Section 2B.
Comparator selection	Single comparator set based on transparent, evidenced criteria.	(1) Narrow-set (2) Economy-wide set	CEPA's approach to select its narrow-set is unclear and appears to be predominantly based on perceived similarity of activities. Furthermore, it contains such a small number of comparators that it is very sensitive to changes in the data (see Annex 4). Contrastingly, its economy-wide set contains a number of industries that are not relevant to gas networks. Our approach uses transparent, well-evidenced criteria, beyond just similarity of activities to select our comparators. It produces a set of comparators that are comparable to gas networks, while including a large enough number of sectors to reduce sensitivity to changes in the underlying data or the volatility of individual observations. We detail our approach in Section 2C.
Mean calculation	Arithmetic	Geometric[176]	An arithmetic mean is more appropriate for shorter time periods than a geometric mean because it is less sensitive to changes in the start date of the series. We detail the full reasoning in Annex 8.

Source: Economic Insight analysis and CEPA's final determination report at RIIO-2[177]

[176] *Note that CEPA does not report whether it uses geometric or arithmetic means. Our replication of its method suggests that it uses geometric means. See Annex 4 for the full replication.*

[177] *'RIIO-GD2 and T2: Cost Assessment –Advice on Frontier Shift policy for Final Determinations.' CEPA (November 2020)*

The table below compares the sectors in our preferred set to those in CEPA's preferred sets for RIIO-2. The key difference between our preferred set and CEPA's narrow-set is that our set contains several manufacturing sectors. This is because CEPA considered manufacturing sectors to not have sufficiently similar activities to gas networks. However, we consider that manufacturing sectors are not objectively different enough from gas networks to exclude them based on this criteria. Furthermore, our approach to comparator selection accounts for a wider range of factors than CEPA's and therefore does not only select sectors only on perceived similarity of activities. Many of the manufacturing sectors are included in our preferred set because they have similar scope for economies of scale to gas networks.

Table 26: Comparison of our preferred comparator set to CEPA's RIIO-2 comparator sets

Our comparator set	CEPA's narrow-set	CEPA's economy-wide set
Construction	Construction	Construction
Wholesale and retail trade; repair of motor vehicles and motorcycles	Wholesale and retail trade; repair of motor vehicles and motorcycles	Wholesale and retail trade; repair of motor vehicles and motorcycles
Transportation and storage	Transportation and storage	Transportation and storage
Total industries	Financial and insurance activities	Financial and insurance activities
Manufacturing		Manufacturing
Chemicals; basic pharmaceutical products		Agriculture, forestry and fishing
Manufacture of rubber and plastic products and other non-metallic mineral products		Mining and quarrying
Computer, electronic, optical products; electrical equipment		Electricity, gas, steam and air conditioning supply
Manufacture of machinery and equipment n.e.c.		Water supply; sewerage, waste management and remediation activities
Manufacture of motor vehicles, trailers, semi-trailers and other transport equipment		Accommodation and food service activities
Manufacture of furniture; jewellery, musical instruments, toys; repair and installation of machinery and equipment		Information and communication
		Professional, scientific and technical activities; administrative and support service activities
		Arts, entertainment and recreation
		Other service activities

Source: Economic Insight analysis and CEPA's final determination report at RIIO-2[178]

[178] *'RIIO-GD2 and T2: Cost Assessment –Advice on Frontier Shift policy for Final Determinations.' CEPA (November 2020)*

11 Annex 6 – The extent to which embodied change is included in TFP

This annex provides further evidence and detail on the arguments we present around embodied technical change in Section 2C.

There has, in the past, been some consensus that productivity metrics such as TFP sufficiently exclude embodied change such that, for the purpose of setting OE, some form of upwards adjustment may be appropriate. For example, in its RIIO-2 Final Determinations, Ofgem rationalised selecting an OE challenge towards the upper end of the supportable range (in part) because it considered the EU KLEMS data sufficiently omitted embodied change.[179] Similarly, in the PR19 Final Determinations, Ofwat raised the issue of embodied change being potentially excluded as a reason to select an OE challenge towards the upper end of its consultant's recommended range.[180]

However, our view is that it is likely that embodied change is included in TFP to some extent and it is likely that the extent to which it is included will vary by industry. We made these assertions in Section 2C and discuss the evidence and intuition underpinning our view below.

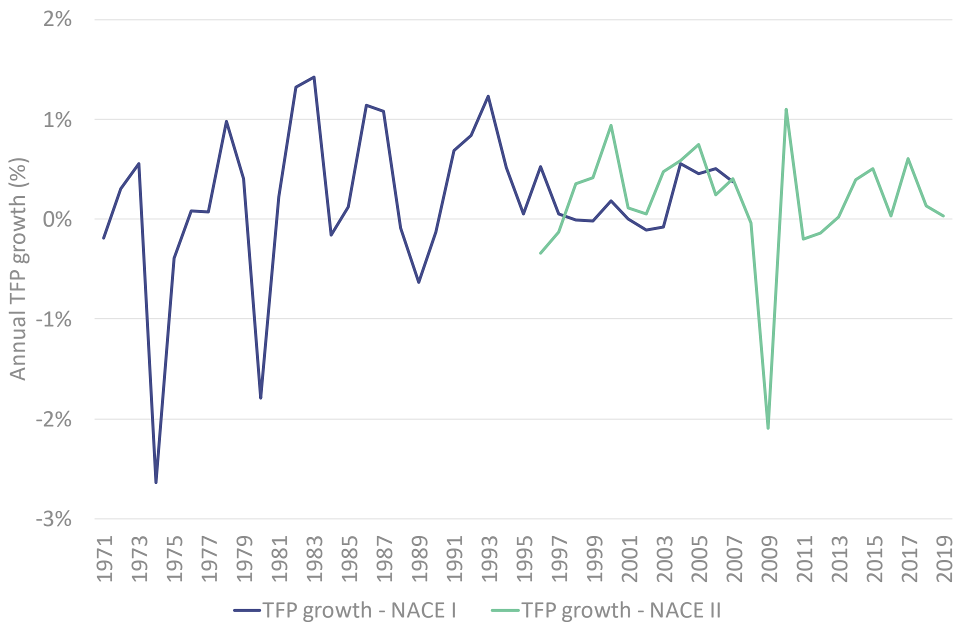
Embodied change is likely included in TFP to some extent

Suppose TFP entirely (or mostly) excluded embodied change. This would mean that the entire (or most of the) variation in TFP over time must be due to variation in 'how well' firms make use of existing technology and assets. This, it is increasingly recognised, seems implausible, especially considering the fact that productivity growth is highly volatile with frequent peaks and troughs (as demonstrated by the figure below). Indeed, if TFP only includes disembodied change (and not embodied change) then the processes used by companies when producing outputs based on their current inputs would need to improve and worsen *substantially*, often within the space of just a few years. We consider this unlikely as it would require companies to suddenly '*go backwards*' after acquiring knowledge and/or establishing best practices and/or making technological improvements relating to their use of *existing inputs*. In other words, it would imply (in extremis) that declines in productivity were due to firms 'forgetting' or 'un-learning' how to use an existing technology / asset as effectively as they once had. It seems more plausible that the large changes in TFP at the UK level are consistent with the data reflecting embodied change to at least some extent.

[179] *Ofgem stated its 'upper end' OE estimate reflected: "an acknowledgment that EU-KLEMS data does not capture cost savings from quality improvements that are embodied in the inputs used by the network companies." See: 'Decision - RIIO-2 Final Determinations - Core Document.' Ofgem (2020); page 50.*

[180] *Specifically, Europe Economics stated: "in order to account for the effects of embodied technical change, a number towards the upper end of each range should be chosen". See 'Real Price Effects and Frontier Shift.' (2018); page 7.*

Figure 22: UK annual TFP growth - EU KLEMS NACE I and NACE II databases



Source: Economic Insight analysis of EU KLEMS data

Indeed, in 2021 (importantly, subsequent to the PR19 and RIIO-2 controls) the Bundesbank made exactly this point as regards to productivity measurement in Germany:

> *"The long-term development of TFP is sometimes also seen as an indicator of disembodied technological progress. In the short term, though, it is difficult to make such an interpretation. Even in the case of severe economic downturns, decreases in technological progress* [from existing technology / assets] *can, if at all, only be regarded to a very limited extent as a plausible explanation for calculated TFP declines. Furthermore, due to its residual character, the contribution of TFP can also pick up other influences on labour productivity. Against this background, there is good reason to interpret TFP more broadly and to view it as a metric of production efficiency."*[181]

Put simply, the Bundesbank is proposing that that technological progress / decline from the utilisation of existing assets (i.e. disembodied change) can only explain a *small* part of the large variations in TFP across Europe. This means that technological progress / decline from the use of new assets (i.e. embodied change) must explain the majority of these variations in TFP.

The above is consistent with the academic literature. In particular, Hulten (2000) makes two key points:

[181] *'The slowdown in euro area productivity growth.' Deutsche Bundesbank Monthly Report (January 2021); page 20.*

- Firstly, the author explains the intuition for why one *might think* that TFP *excludes* embodied change by referring to the Solow paradox. This relates to various arguments made that the slowdown in productivity observed in the USA throughout the late 1960s and early 1970s, is inconsistent with the benefits that are believed to have occurred from the technological change at the time (such as the computing revolution). In relation to this, Robert Solow in 1987 famously argued that: *"[y]ou can see the computer age everywhere but in the productivity statistics."*[182] Hulten summarises this point more broadly as: *"one might well say that we see new technology everywhere but in the productivity statistics."*[183]

- Secondly, Hulten explains the counterargument i.e. why one *might think* embodied change is *included* in TFP. The author states that "*there is another 'new economy' paradox that has been largely overlooked: if the missed quality change* [arising from new technology] *is of the magnitude suggested above* [an upward bias of 0.6 percentage points in CPI per annum to account for quality improvements], *the quality of the goods in past centuries – and the implied standard of living – must have been much lower than implied by official (and allegedly quality-based) statistics. Indeed, taken to its logical conclusion… quality adjusted average income in 1774…* [would be] *dubiously small."*[184] In essence, the author is suggesting that growth in average income has been so significant over time that, if embodied change is completely excluded, then the starting point for average income (and implied living standards) seems implausibly low.

Hulten then summarises the above two points by concluding that that: *"In other words, conventional estimates of productivity growth are either much too large or much too small, depending on one's view of the matter. The truth undoubtably lies somewhere between the two extremes."*[185]

Both EU KLEMS documentation and a response to our query by the ONS are also consistent with the above (i.e. that TFP will include embodied change to *some* extent). In particular:

- An overview of the EU KLEMS dataset states the following: *"[u]nder strict neo-classical assumptions, TFP growth measures disembodied technological change. In practice, TFP <u>is derived as a residual</u> and includes a host of effects such as improvements in allocative and technical efficiency, changes in returns to scale and mark-ups <u>and technological change proper</u>. All these effects can be broadly summarised as "improvements in efficiency", as they improve the productivity with which inputs are being used in the production process. In addition, being a residual measure, TFP growth also includes measurement errors and the effects from unmeasured output and inputs".*[186] In summary, this suggests that the EU KLEMS TFP (and ONS MFP) *mainly* but, crucially, *not exclusively* reflects disembodied change.

182 *'Total factor productivity: a short biography.' Hulten, C; NBER (January 2000); page 2.*
183 *'Total factor productivity: a short biography.' Hulten, C; NBER (January 2000); page 3.*
184 *'Total factor productivity: a short biography.' Hulten, C; NBER (January 2000); page 3.*
185 *'Total factor productivity: a short biography.' Hulten, C; NBER (January 2000); page 4.*
186 *'An overview of the EU KLEMS Growth and Productivity Accounts.' European Commission (October 2007).*

- The ONS told us that: *"whilst multifactor productivity should measure just the disembodied change, we do think that <u>there is likely some embodied change in the measure</u>."* This was in response to our query to the ONS on the appropriate interpretation of the data (noting that EU KLEMS draws on ONS data).

In summary, the exact *proportion* of embodied change included in TFP is unclear, but the evidence suggests that it is included to *some degree.*

The extent to which TFP includes embodied change likely varies by industry

GO measures of TFP include inputs from labour, capital and intermediate inputs (while value added measures use just labour and capital). Quality adjustments are then applied to capital assets and intermediate inputs in an attempt to remove the impact of quality changes over time.[187] If these quality adjustments were perfect then TFP would only measure disembodied change. However, it is likely that: (i) there is a margin of error in this approach; and (ii) this margin of error varies by industry because some industries use relatively more capital and intermediate inputs (which we show in Section 4B). Therefore, the extent to which all of the 'quality' can be removed varies by industry with those industries that use relatively more capital (and intermediate inputs) being more likely to include a greater proportion of embodied change in their TFP data.

Illustrative example of embodied technical change

We set out an illustrative example of how the adjustments implied by Table 6 could be calculated in theory. In practice, the information presented below would not be available as it is unknowable how much embodied change is included in TFP estimates.

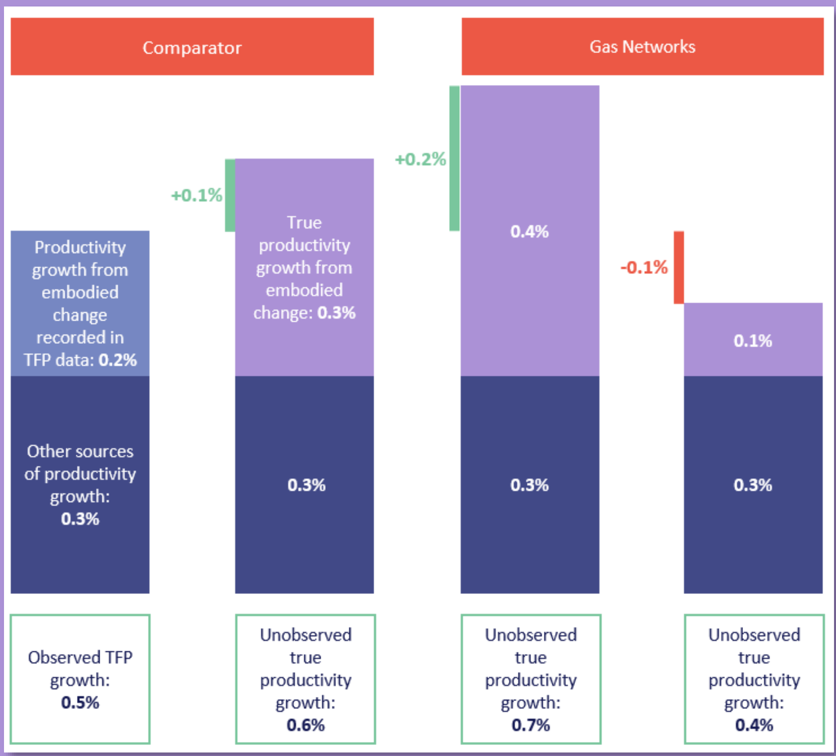
[187] *The quality adjustment is not applied to labour in the same way. Specifically, the labour input measure in the ONS MFP dataset is called the "Quality-adjusted labour input". However, we note that the quality adjustment is different for labour as it aims to adjust the number of hours worked by employees, based on the following factors that represent their "quality": education level, industry of employment, age and gender. Therefore it would be necessary to adjust the mix of employees for this "quality" to increase over time. This is in contrast to embodied change, which relates to changes in the quality of the <u>same</u> inputs over time i.e. without changing the mix of inputs used.*

Box 1: Illustrative example of embodied technical change

Suppose a *comparator* industry could achieve embodied technological change related productivity gains of 0.3% pa, but only 0.2% of that was included in its TFP data (and suppose its annual average TFP was 0.5%). At face value, this would imply that, in inferring an OE challenge for gas networks from the comparator TFP data, one would need to make an upwards adjustment of 0.1% for (excluded) embodied technological change (i.e. 0.3% of embodied change related gains, less the 0.2% of that included in the TFP data = 0.1%). However, this is not necessarily correct. For example:

If, in fact, the *gas networks* could achieve higher embodied technological change related gains than the comparator of (say) 0.4% pa, the above adjustment would be too small (i.e. OE will still be under-stated, even having made the adjustment). Instead, an upward adjustment of 0.2% would be appropriate.

Alternatively, if the *gas networks* could achieve lower embodied technological change related gains than the comparator of (say) 0.1%, the above adjustment would be too large (i.e. OE will be overstated, having made the adjustment). Instead, a downward adjustment of 0.1% would be appropriate.

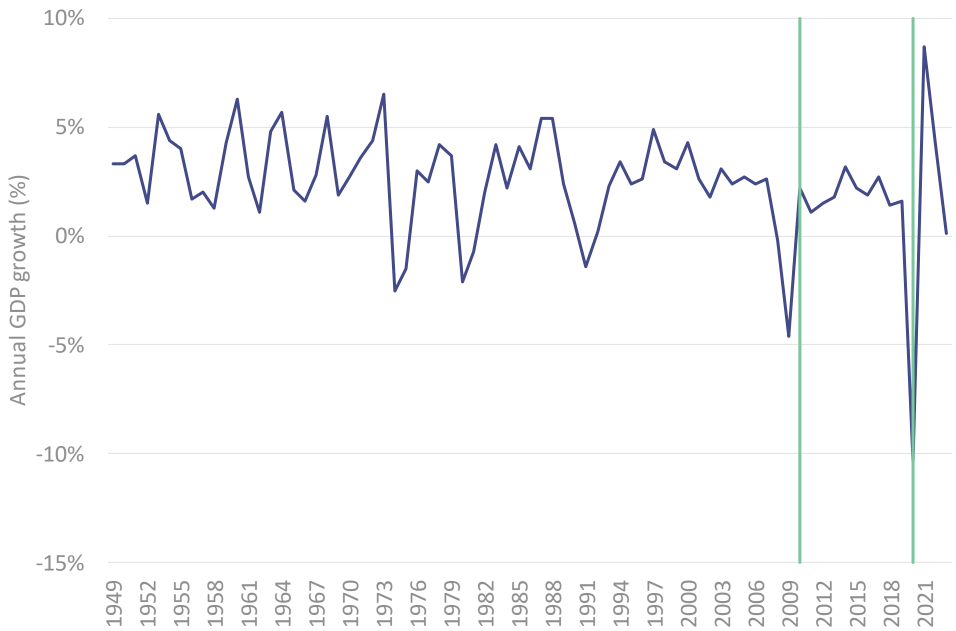


12 Annex 7 – Business cycle analysis

In Section 2B we note that 2010-2020 is the most recent business cycle. This annex details why this is a business cycle, as well as outlining when any other relevant business cycles occur.

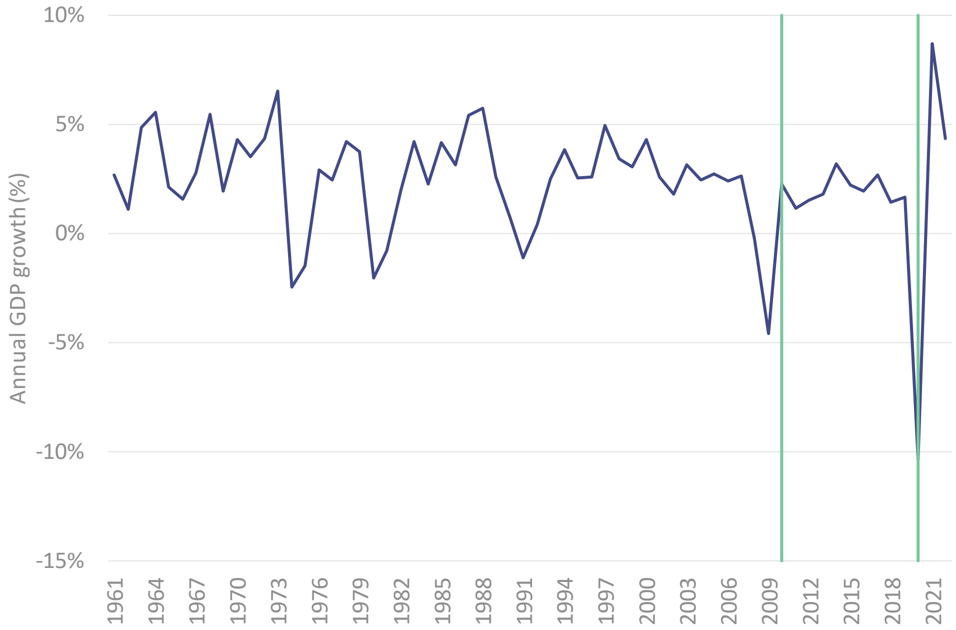
The figures below present annual data on UK GDP growth, published by the ONS (covering the period 1949-2023) and the World Bank (covering the period 1961-2022) respectively. A trough-to-trough analysis of these figures (shown by the green, vertical lines) indicates that the most recent business cycle is 2010-2020 (with 2010 corresponding to the year following the end of the previous economic cycle in 2009).

Figure 23: Recent business cycle based upon ONS annual GDP growth



Source: Economic Insight analysis of ONS data

Figure 24: Recent business cycles based upon World Bank annual GDP growth



Source: Economic Insight analysis of World Bank data

This trough-to-trough analysis is broadly consistent with peaks and troughs published by the Economic Cycle research Institute (ECRI). The ECRI have published peak and trough dates for business cycles across 22 different countries (including the UK) since the 1970s. The table below reports these, and indicates that the most recent business cycle now corresponds to 2010-2020 (consistent with our analysis of ONS and World Bank data above).

Table 27: UK business cycle peak and trough dates from ECRI (1974 – 2020)

Business Cycle	Peak / Trough	Dates
1974-1975	Peak	September 1974
	Trough	August 1975
1975-1981	Peak	June 1979
	Trough	May 1981
1981-1992	Peak	May 1990
	Trough	March 1992
1992-2010	Peak	August 2008
	Trough	January 2010
2010-2020	**Peak**	**October 2019**
	Trough	**April 2020**

Source: Economic Insight analysis of 'Business Cycle Peak and Trough Dates, 22 Countries, 1948-2020.' (ECRI) (last accessed 9 February 2023).

13 Annex 8 – Geometric vs arithmetic mean

We note in Section 2C that we use arithmetic means to calculate the average growth rate of TFP across the time period under consideration, for each of our comparator sectors. There are two different methodologies that can be used to calculate the average growth rate in TFP (or MFP) over time. These are as follows:

- **Arithmetic mean.** This is a simple average that we calculate by taking the sum of a series of numbers, and dividing this sum by the count of that series of numbers (i.e. the number of numbers in the series).

- **Geometric mean.** We calculate this by taking the product of a series of numbers, raised to a power equal to the inverse of the length of that series (i.e. the number of numbers in the series).

Each of these methods have advantages and disadvantages, which we summarise as follows:

- **Geometric means are more appropriate for series that are not independent.** Arithmetic means are appropriate when the values in a series are independent of one another; but when the values are not independent will produce less reliable results. In contrast, geometric means can be used reliably when the values in a series are not independent from one another. This allows geometric means to account for the influence of a variation in a single year on any future years, which is a common feature of economic data.

- **Geometric means suffer from greater sensitivity to the start date over which the average is taken compared to arithmetic means.** This is because geometric means compound values based on the start value of the series, whereas the arithmetic calculation does not use compounding.

Based on the advantages and disadvantages of each approach, our view is that arithmetic means should be used to assess productivity over a shorter time period (less than about 10 years), while geometric means should be used over a longer time period (more than about 10 years). However, this threshold should be taken as a guide, rather than a strict rule. Our view is based on the following:

- Over a short period of time, the average productivity growth will be sensitive to the start data; and it is beneficial to assess the impact of year-on-year volatility. This suggests an **arithmetic mean** is more suitable for short time periods.

- Over a longer period of time, it is beneficial to strip out the year-on-year volatility to determine the actual long run productivity growth rate. In addition, the calculation will be less sensitive to small variations in the start date over a long period of time. This suggests that a **geometric mean** is more appropriate for longer time periods.

We use an arithmetic mean to calculate all the results in the main report (as we note in Section 2C) and all the results in Annex 1 and Annex 3. Outside of this annex, we only use geometric means when replicating and updating CEPA's RIIO-2 method in Annex 4. In order to allow a comparison between the two methods, we present the results of our *'recommended range'* using both the geometric and arithmetic mean in the table below. As can be seen in the table below, the estimates using the geometric mean are either identical or only fractionally different to the estimates using the arithmetic mean.

Table 28: Comparison between geometric and arithmetic mean

Comparator	2010-2019		1970-2019	
	Arithmetic	Geometric	Arithmetic	Geometric
Final results (average)	0.2%	0.2%	0.8%	0.8%
Manufacturing	0.4%	0.4%	0.9%	0.9%
Chemicals; basic pharmaceutical products	1.2%	1.1%	1.6%	1.5%
Manufacture of rubber and plastic products and other non-metallic mineral products	1.0%	1.0%	0.9%	0.9%
Computer, electronic, optical products; electrical equipment	1.3%	1.3%	2.2%	2.2%
Manufacture of machinery and equipment n.e.c.	-0.2%	-0.3%	0.9%	0.9%
Manufacture of motor vehicles, trailers, semi-trailers and of other transport equipment	-0.1%	-0.2%	0.7%	0.7%
Manufacture of furniture; jewellery, musical instruments, toys; repair and installation of machinery and equipment	-0.4%	-0.5%	1.0%	1.0%
Construction	-0.1%	-0.1%	0.1%	0.1%
Wholesale and retail trade; repair of motor vehicles and motorcycles	-0.1%	-0.1%	-0.1%	-0.1%
Transportation and storage	-0.6%	-0.7%	0.5%	0.5%
Total industries (A-S)	0.2%	0.2%	0.2%	0.2%

Source: Economic Insight analysis of EU KLEMS data

14 Annex 9 – Survey methodology

This annex provides a summary of the method used to conduct the survey of leading academic experts on UK productivity growth. As we note in Section 3A, this survey was undertaken as part of a wider research exercise on UK productivity, around which a working paper has been developed[188] that provides a more detailed description of the survey than the summary below.

The survey was distributed to academic experts on productivity analysis via email. It was completed by 26 of the experts, and asked them to provide their views on:

(i) The factors driving the lower level of productivity growth observed in the UK since 2008, and the reasons behind these factors.

(ii) The factors driving historical differences in productivity growth between sectors, and the reasons behind these factors.

(iii) The prospects for UK productivity growth in the future. In particular the experts were asked to provide their views, and the reasons for their views, over three separate time periods, the next: (a) twelve months (2024); (b) five calendar years (2024-2028); and (c) ten calendar years (2024-2033).

(iv) Which sectors are expected to out- or under-perform the UK average in the future and their reasoning for this.

[188] *'The UK productivity puzzle: A survey of the literature and expert views.' Williams, S.; Glass, A.; Matos, M.; Elder, T.; and Arnett, D. (January 2024).*

15 Annex 10 – Quality gains in TFP

In the note overleaf, Professor Anthony Glass (an expert in productivity analysis) explains why a 'double-count' of efficiency gains may occur if companies are simultaneously set an OE challenge for costs and expected to make quality improvements.

A Note on:

Why the TFP metric for comparator industries captures productivity gains arising from increased output / quality?

Anthony Glass*,#

May 2024

Total factor productivity (TFP) is simply a measure of the change in output (or value created if a strict quantity-based output measure is not used, which is typically the case) that cannot be explained by a change in the quantities of all the inputs (say labour, capital and energy). In other words, in the case of, for example, a particular industry (denoted below as the *ith* one) in year *t*, TFP_{it} measures how effectively the industry is at transforming the quantities of its inputs into output. By decomposing TFP_{it} into its components, one can assess the sources of improvement and declines in TFP_{it}. TFP_{it} has the following four theoretically founded components.

$$TFP_{it} = TC_{it} + RTS_{it} + TE_{it} + AE_{it}, \tag{1}$$

where TC_{it} is technical change, RTS_{it} is returns to scale, TE_{it} is technical or economic efficiency, and AE_{it} is allocative efficiency. Note that in domestic public utility industries TC_{it} is often referred to as the frontier shift, FS_{it}.

One could obtain an historical estimate of annual FS_{it} from a fitted cost model for public utility companies, which is then used to set an annual frontier improvement target for companies for the next 5-year regulatory period. Domestic public utility regulators, however, directly set the annual frontier improvement targets for the companies based on directly available TFP data for comparator industries from databases such as EU KLEMS and from the ONS. Since from Eq. 1 FS_{it} is one of the four components of TFP_{it}, the annual frontier improvement targets for companies that domestic public utility regulators set based on the TFP data of comparator industries will be an overstatement. Turning now to the interpretation of this directly available TFP data.

Crucial to the interpretation of the TFP measures of comparator industries is the appreciation that they are ordinarily based on two output measures – gross output (GDP) and value added. Consider the case of value added, which is net of intermediate outputs, so it is simpler, therefore, to appreciate how increases in the value added-based TFP measures of comparator industries represent gains from increased output and / or quality. That is, value added reflects the value generated by an entity's own processes, as it represents the net output of the entity after subtracting intermediate inputs from gross output. A value added measure does not distinguish between whether an output gain reflects 'reductions in production costs' and / or 'revenue increases', where the latter may be due to an increase in quantity and / or price. Based on companies in the comparator industries operating in competitive markets, so that quality flows through to prices, we can interpret a value added-based output gain that is due to a revenue increase as reflecting an increase in the quantity supplied and / or an increase in the quality of the output.

Whilst we considered the case of the value added of comparator industries to simplify matters, an increase in gross output (or any monetary measure of output for that matter) of a comparator industry will typically reflect some combination of a reduction in production costs and an increase in quantity and / or quality. Relating this back to the setting of the annual frontier improvement targets for companies by domestic public utility regulators, the implication is that if the entirety of the annual frontier improvement targets are allocated to cost savings, but companies are simultaneously required to make quality improvements and /or possibly increase the quantity of at least one output, there will be a 'double-count'. Consequently, the size of the ongoing productivity frontier improvement that companies would be required to make would be greater than the implied benchmark for this improvement from the comparator industries.

* Professor of Managerial Economics, Sheffield University Management School (SUMS), Conduit Road, Sheffield, S10 1FL.

The views expressed in this note are those of the author and do not necessarily reflect those of the University of Sheffield.

WE MAKE ECONOMICS RELEVANT

Economic Insight Limited

125 Old Broad Street
London
EC2N 1AR
+44 207 100 3746
www.economic-insight.com