Article

# A guide to quality adjustment in public service productivity measures

This guide explains what a quality adjustment is, including important underlying concepts, statistical properties, and application in the National Statistic.

Contact:
Leah Harris
productivity@ons.gov.uk
+44 (0)1633 456525

Release date:
7 August 2019

Next release:
To be announced

# Table of contents

# 1 . Overview

This guide explains the concept and implementation of quality adjustments in the measurement of public service productivity. A quality adjustment is, in its simplest terms, a statistical estimate of the change in the quality of a public service. Whilst we aim to achieve complete coverage of the public services, currently these are only applied in four service areas: Healthcare, Education, Public Order and Safety, and Adult Social Care. Following international guidance, quality adjustments are not applied to public sector output measures in the National Accounts, only to the Public Service Productivity series.

Sections 2 and 3 cover the principles of adjusting for quality, developed using some important underpinning concepts. One such concept is attributability. This is the question of, how far can we attribute an observed outcome to the provision of a service? For example, the outcome may be good grades at school; to what extent did the school contribute to the achievement of these grades?

Section 4 applies these concepts in the context of Public Order and Safety – a service area we have recently developed a quality adjustment for in our annual estimates. This follows a five-step framework that can be summarised as follows:

- Define dimensions of quality for the service area.

- Decide what can be measured.

- Consider the limitations of each quality indicator.

- Design the quality indicator indices.

- Process and aggregate the data.

The history of quality adjustment in the UK spans two decades at the Office for National Statistics (ONS), including the Atkinson Review: Final Report (PDF, 1.07MB), published in 2005. More information on the history of quality adjustment and associated work at the ONS is in Section 6 of this guide.

Because of this work, the ONS is a world leader in measuring public service productivity, and Section 7 surveys work by the international community in this field, offering interesting extra reading and ideas.

# 2 . What is quality adjustment?

Public service productivity is how much output is produced per unit of inputs, for public service areas. Public services are services provided to users which are provided by a public body, or purchased by a public body for use. The service is free, or has only nominal charges, at the point of delivery.

It is important to measure as the public sector forms around a fifth of UK gross domestic product (GDP), but presents some unique challenges in doing so. One of these is how to account for changes in the quality of the service provided. For a normal market good, an increase in price would normally reflect an improvement in quality, but as public services have no market price, we cannot use prices to assess these changes. Whilst higher-cost public services are often of greater value than lower-cost services (such as heart surgery relative to a dental check-up), using cost alone may not sufficiently differentiate between high and low value services. As such we look to identify direct measures of output in volume terms and, where possible, adjust these for their impact on the quality of outcomes achieved. For an overview of the challenges in measuring public service productivity, see our blog post. Important for understanding quality adjustment is the relationship between output and outcomes. The adjustments account for the outcome of the service rather than just the output. Output is what a public service provides, and the outcome is the end effect of the service.

To illustrate how factoring in these quality changes works, we can use the example of students at school. More students at a school means higher output, which would increase productivity (if inputs remain constant). However, if these students are achieving lower grades in their exams, this could be an indication of falling quality. The quality adjusted productivity index will account for this and will either grow at a slower pace or fall (again, assuming no changes to inputs growth). A comparison of the total productivity index with quality adjustment and without is in Section 5. The quality adjustment can show how successful the service is at achieving these end goals through what it provides.

Of the nine service areas included in our annual estimates of total public service productivity, four are quality adjusted in some way – see Section 5 of this guide for more detail, and Section 6 of the Quality and Methodology Information report. They are Healthcare, Education, Public Order and Safety and Adult Social Care.

Public services can be grouped by whether they are collective or individual. Collective services, such as Defence, are those provided to society as a whole rather than individuals. It is difficult to identify and measure output directly in the collective services. As such, we generally use, as a last resort, the "output-equals-inputs" convention, meaning they typically have a productivity growth rate of zero. Individual services are those provided directly to individual people or households, such as medical operations.

Direct measurement of output is easier in individual services. It is desirable to quality adjust collective services, but currently only individual services are adjusted for quality in our public service productivity statistics.

It can be difficult to devise an adjustment that accurately measures outcomes; importantly, how attributable to the service provided is a successful outcome? A reduction in crime could stem from the number of police officers, but it could also stem from a multitude of other sources. These could include reduced inequalities, higher life satisfaction, better mental health services, or the implementation of social activities and programmes within the local community. This topic and others are discussed further in Section 3 of this guide.

# 3 . What makes a good quality adjustment measure?

## The conceptual foundation

What benefits does someone receive from being provided a public service? How much of that benefit is directly because of the service?

These are the first two questions that should be asked when thinking about quality adjustment.

For example, in Education, the number of students attending school is used as a quantity output measure, and their attainment at around age 16 (GCSEs or Scottish Nationals, for example) is used as a quality measure. Attainment of these students may serve as a suitable indicator for changes in the quality of the education they receive. But, looking at the two important questions raised above: firstly, is attainment the only (or most important) benefit that students get from attending school, and secondly, how much of their attainment is attributed to the school's services?

The extent of the attribution of an outcome to a public service is one of the most important ideas to consider when adjusting for quality. Quality adjusting is used to show the impact on successful outcomes from the service provided; it aims to show how the output provided in this service area affects the outcome (end goal) of the service. We need to include all desirable outcomes that the service is expected to affect. This is unlikely to be achievable, but we can develop a good approximation of quality growth and its effect on output, and productivity.

Returning to our 16-year-old students, the literature[1] would suggest that their social skills and mental well-being are also affected significantly by their education, and that factors such as family, home environment, and personal motivation and interests have a significant impact on their grades at school. Ideally, quality adjustments should incorporate this evidence in some way. Alternatively, the limitations of the adjustment should be known, and alternative options or methods should be consulted on with stakeholders, when required.

In Section 2, the relationship between output and outcomes was described. An important feature of this relationship is that there will likely be a significant time lag between the provision of a public service and the possible observation of its full impact on an individual. For example, a good primary education will positively benefit the students long after they leave primary school. This time lag effect will make observing the contribution of the output to outcomes more difficult, as there will be a wait for full data. The longer the observation period, the more complex the potentially confounding network of factors affecting the outcomes becomes. From the point of view of statistics users, a longer wait between publication of data and the years the data cover is less desirable; the Bean Review (PDF, 5.1MB) emphasised the need for timelier statistics. For public service productivity, quality adjustment is currently only used in our annual National Statistic, which has a two-year time lag. Our experimental quarterly series includes no quality adjustment, but provides the timelier information that some users require.

## Methodological and statistical properties

Several pieces of research have investigated criteria that should be used to assess the suitability of a quality adjustment. The Centre for Health Economics (CHE) at the University of York built on earlier work of the US Agency for Healthcare Research and Quality, and UK National Centre for Health Outcomes Development and Health and Social Care Information Centre, to derive a set of criteria. With reference to this earlier work, in particular the CHE's Accounting for the Quality of NHS Output (PDF, 2.2MB), the following criteria can be used to assess quality adjustments:

- Is the adjustment relevant and does it have an appropriate degree of coverage?

- Is the adjustment easily measurable?

- Is the adjustment and the data it uses representative of the entire population?

- Are the required data of high quality and reliable?

- Is there an acceptable level of assumptions made in the methodology used to produce the adjustment?

- Are the required data released at appropriate intervals?

- Are the required data sustainable; that is, will it continue to be released?

- What is the time lag between the current period and the release date of the associated data?

- How sensitive is the adjustment to change and how likely are revisions to the data?

- What are the errors in the estimate like? Are they systematic in one direction, or changing in size over time?

To account for quality changes in different parts or components of a service, it is useful to differentiate between them and measure them separately when possible. This was emphasised in the Atkinson Review. For example, in Public Order and Safety (POS), output is measured by fire protection services, probation, prisons, and courts (which is split into five further subcomponents). More specific quality adjustments can then be used, such as courts' timeliness for some courts components only, and to identify the cause of changes in overall POS productivity growth.

From this idea, the relative importance of different aspects of quality should be accounted for. For the overall quality adjusted index in a service area, each adjustment is weighted, with a view to weight more valuable adjustments more highly. Of course, this introduces another area of investigation; which adjustments add the most value?

## Meeting the needs of stakeholders

A good quality adjustment should also consider the needs of the statistic's users. How useful would the measure be to the statistic's stakeholders, and who specifically may be interested in the result of the adjustment? For example, an updated primary schools adjustment will be of interest to the Department for Education, and possibly to educational charities and think tanks. Close collaboration with these parties allows for important understanding of service area specific techniques and ideas, such as the use of quality adjusted life years (QALYs) for Healthcare.

Finally, the quality adjustment's suitability should be monitored over time. It is recommended in The Atkinson Review that triangulation should be used to assess different data sources and methods as part of reviews of the current methods, evidenced by the Office for National Statistics (ONS) in Quality adjustment for public service education: triangulation (PDF, 190.6KB). This is essential to ensure that the estimates that are produced remain of a high quality.

As public service productivity measures activity in the public sector, changes to government policy could affect the quality adjustment. For example, for the Education quality adjustment in England, the average point score system for secondary school attainment was replaced by the proportion of students receiving a certain number and grade of qualifications (five A* to C grades at GCSE). This change came into effect from 2013 (PDF, 183KB), because of reforms of the education system and which qualifications were included in school performance tables. Recent policy changes to GCSE curriculums and grading systems have led to the Department for Education now providing "Attainment 8"[2] as their headline attainment measure. This illustrates that quality adjustment measures need frequent review and improvement, using the ideas described in this section. Working with relevant government departments and public bodies is an important part of this process.

The principles described here are applied in Section 4, where the process of developing a quality adjustment is described with reference to a specific service area, Public Order and Safety, to bring the ideas into a practical light.

### Notes for: What makes a good quality adjustment measure?

1. Two examples are What makes a test score? The respective contributions of pupils, schools, and peers in achievement in English primary education (PDF, 747KB) and The link between pupil health and wellbeing and attainment (PDF, 167KB).

2. Attainment 8 is a new scoring system that aggregates students' marks across eight subjects – see the Department for Education's Secondary accountability measures guide (PDF, 2.1MB) for more details.

# 4 . Designing a quality adjustment – in the context of Public Order and Safety

This section covers the main steps towards designing and implementing an adjustment, using Public Order and Safety (POS) as an example. This was the process behind the recent development of quality adjustment for POS, as described in Quality adjustment of public service criminal justice system output: experimental method in 2017, and in Quality adjustment of public service public order and safety output: current method, which was published after the POS adjustment was approved for incorporation into the National Statistic in 2018.

The following steps show that the process of designing a quality adjustment index should be creative and inclusive of all outcomes, but mindful of the limitations of the chosen measures and of what the data really show. To that end, the process could be summarised as follows:

1. Define dimensions of quality for the service area: Think about what the service does and what its desirable outcomes are. How might fulfilling these outcomes be expressed with data? How attributable to the service are these desirable outcomes?

2. Decide what can be measured: In the light of the outcomes and potential measures identified, what data are actually available? Do these meet the essential statistical and theoretical criteria?

3. Consider the limitations of each quality indicator: Which limitations or caveats are there with each one? Which limitations can be accounted for by including other adjustments or by adapting the source data? Which data sources fall under acceptable error margins, and therefore should be dismissed?

4. Design the quality indicator indices: Regularly check that the index meets the previously identified criteria for suitable use and that the data represent quality change in the appropriate dimension. Try several different methods to find the best, if necessary, with different adjustments and weights in each, for example. Use triangulation evidence to assess alternative options involving consulting with stakeholders and sectoral experts.

5. Process and aggregate the data: First, the adjustment or adjustments for each output component, then for the combined output series, and finally include in the productivity calculations.

These steps are explored in more detail in the rest of this section.

Firstly, it is necessary to define the dimensions of quality for the service area. These are indicators of the service meeting the needs of its users, government, and society. Each indicator should be theoretically sound, with literature available to support the argument for its inclusion. Returning to the list of statistical criteria in Section 3, we need a reliable and timely data series for the indicator. Finally, it is desirable to account for multiple dimensions of quality, which may require multiple data sources. If so, it is important to consider the interconnectedness of the various quality measures and weight them accordingly in the overall quality adjustment.

To identify the dimensions of quality, the desirable outcomes for the service should be used. These can be identified for POS by using the Ministry of Justice's strategic objectives, published in formats such as the single departmental plan 2019 – 2022. Here, there are eight objectives, including "Provide a transparent and efficient court system" and "Reduce rates of reoffending and improve life chances for offenders". These objectives demonstrate the successful outcomes of POS and quality indicators that could be used. For example, an efficient court system could be reflected by data on the speed with which they process cases. The indicators may be considered representative of an outcome that is affected by the whole service, or they may only apply to one or several subcomponents of POS output. The courts' timeliness measure only applies to courts output.

Note that Police is a separate service area within public service productivity statistics and as such this quality indicator will not reflect police service provision.

The subcomponents for POS output are shown in the table below, alongside the five year average of their expenditure shares from 2012 to 2016 to indicate their relative size:

Table 1: Subcomponents for Public Order and Safety output and respective five-year average percentage shares of Public Order and Safety expenditure, 2012 to 2016

| Sub-component | Percentage share of 2012 to 2016 expenditure (five-year average) |
|---|---|
| Prisons | 29% |
| Fire | 22% |
| Legal Aid | 13% |
| Probation | 9% |
| Crown Prosecution Service | 9% |
| Magistrates and other LA Courts | 7% |
| Crown Courts | 6% |
| County Courts | 6% |

Source: Office for National Statistics

Notes

1. There is one other component called "Public order and safety not elsewhere classified" which is measured in the Other service area in the Total public service productivity estimates. In the Other service area output is assumed to be equal to inputs. Back to table

2. Crown Courts, County Courts, Magistrates and Other Local Authority Courts, Crown Prosecution Service and Legal Aid are all combined into Courts in the National Statistic article, where analysed. Back to table

3. Shares may not sum to 100% due to rounding. Back to table

The POS adjustment uses an index of reoffending rates to show the effectiveness of the justice system at rehabilitating offenders, applied to prisons, probation, and most of courts output. However, it is not applied to the entire POS output index – fire protection services has a different set of desirable outcomes. County courts, which process civil rather than criminal cases, are also excluded from the quality adjustment, whereas Magistrates and Crown Courts, which predominantly process criminal rather than civil cases, are included.

For prisons specifically, quality indicators could include: physical and emotional well-being within prisons for staff and inmates; the escape rate; or the activities that inmates can partake in. For courts, they could include the appeal rates against convictions, or the speed at which they process a case. These options were reviewed in a discussion article in 2017, and mostly found to be unsuitable or infeasible. For example, data on the types of accredited courses completed by inmates were considered. It was rejected because course completion data were not to a high standard.

Once there is a list of potential indicators of quality, there must be an assessment of the data available for inclusion in the final adjustment. Importantly, data should not be used simply because they are available and might indicate a quality change; available data must meet the list of criteria in Section 3 and must closely follow the theoretical and conceptual evidence. Making unreliable links between a data series and what it actually means should be avoided. This is why it is important to use a range of indicators for different dimensions of quality and apply them to as specific a subcomponent or group of components as possible.

The quality indicators used in our POS adjustment are shown in the table below, alongside their data sources:

Table 2: Quality adjustments for Public Order and Safety and their data sources

| Quality adjustment indicator | Sub-components it is applied to | Data source |
| --- | --- | --- |
| Recidivism (severity adjusted re-offending rates) | Prisons; Probation; Magistrates Courts; Crown Courts; Crown Prosecution Service; Legal Aid | Ministry of Justice, Office for National Statistics |
| Custody escapes | Prisons | Ministry of Justice |
| Safety in prisons | Prisons | Ministry of Justice, Health and Safety Executive |
| Courts' timeliness | Crown Courts; Magistrates Courts | Ministry of Justice |

Source: Office for National Statistics

These adjustments were chosen for use in the final index in part due to the wide range of appropriate data available from the Ministry of Justice.

Ideally the data underlying any quality adjustment should be consistent over time and be available throughout the full time series. However, a compromise between this and data availability may be necessary. In POS, recidivism is applied from 2000 onwards, prisons safety and custody escapes from 1997, and courts' timeliness from 2011.

Recidivism is applied to all subcomponents (except fire protection services and county courts, neither of which are presently adjusted for quality). Recidivism is the reoffending rate, weighted by the severity of the crime committed for the reoffence, and by the characteristics of the offender. This helps control for other factors that affect reoffending rates and so helps isolate the effect of the quality of the justice system. The other adjustments are separated out across certain subcomponents. More detail will be provided later in this section.

There will most likely be limitations for any adjustment that is chosen; there will be a compromise between how much uncertainty is permissible against how useful the adjustment is. Adjusting for quality in some form is essential to accurately measure the productivity of public services, as demonstrated in the Atkinson Review and a range of other literature[1] .Therefore, it is accepted that no adjustment is perfect and the best available indicator of a dimension of quality is acceptable. We have a constant review and development process in place to assess our adjustments and how they can be improved, or how we can develop further indicators for service areas that currently are not quality adjusted.

For POS adjustments, consider recidivism again. We account for the characteristics of the offender, using the Offender Group Reconviction Scale[2] (OGRS4/G), from 2005 onwards. It uses age, gender and criminal history to assess the reoffending risk of a given group of offenders by producing a score between 0 and 1, and is based on extensive research on how different characteristics can affect reoffence rates. The Ministry of Justice has conducted various studies on this – for example, see Do offender characteristics affect the impact of short custodial sentences and court orders on reoffending? (PDF, 400KB) .

Our recidivism adjustment is also constructed to consider how severe the reoffence is; a severity weighting for the reoffender's crime is applied. It is derived from our Crime Severity Score for England and Wales. These weights are shown in the table below:

Table 3: Severity weights for different offences

| ONS Crime Group | Implied Severity |
|---|---|
| Sexual offences | 24 |
| Robbery | 12 |
| Violence against the person | 2.2 |
| Fraud | 1.4 |
| Theft offences | 1.3 |
| Criminal damage and arson | 1 |
| Other crimes against society | 1 |
| Summary[1] | 0.5 |

Source: Office for National Statistics

Notes

1. Summary crimes includes less serious crimes such as television licence evasion which are not captured by the Crime Severity Score and are given a score of 0.5, half the value of the next lowest score of 1.0. Back to table

2. Weights are relative to a baseline of 1.0 for Other crimes against society. Back to table

3. Some categories cover a wide range of crimes – Violence against the person covers all from minor affray offences to murder, which are weighted by volume to give the final Implied Severity measure. The large number of minor crimes in this group are why its weighting is low compared to Sexual offences and Robbery. Back to table

The fact that reoffending is affected by offender characteristics, and that different offences have different implications for the victim of the crime, are two examples of a limitation or caveat that must be considered. The design of the quality adjustment needs to account for these.

A limitation of using custody escapes is that the data only show a small number of escapes each year. Therefore, an increase of only two or three escapes will appear as a large change, despite forming a very small proportion of the total prison population. To account for this, a former Ministry of Justice key performance indicator is used; a baseline of 0.5% of the prison population. The difference between this baseline and the number of escapes is compared, which reduces the volatility of the series.

The prisons safety adjustment uses data on self-inflicted injuries and fatalities, and injuries and fatalities that occur through assault from another person, for both prison staff and inmates. The injuries or fatalities are grouped into one of three categories; "Less severe", "Severe" and "Those resulting in a death". The number of incidences are aggregated. However, to reflect the differing severity of these incidents, each of the three categories are weighted differently. For instance, a "Less severe" injury is less indictive of a safeguarding failure than a death. Weightings from the Health and Safety Executive's (HSE) costs to society of workplace injuries are used. This is the best approximation of costs of a prison incident available, as the HSE considers human and financial costs. However, it is not specific to inmates and as such the weighting system for prisons safety is under review.

Now that the indices for each indicator of quality have been constructed, they need to be weighted together, to produce the quality index for each subcomponent. The weights used for POS adjustments are presented in the table below.

Table 4: Quality adjustments for Public Order and Safety and weighting of each adjustment for each subcomponent

| Component | Recidivism | Prison safety | Custody escapes | Courts' timeliness |
|---|---|---|---|---|
| Prisons | 29.2% | 37.5% | 33.3% | |
| Probation | 100% | | | |
| Magistrates Courts | 50% | | | 50% |
| Crown Courts | 50% | | | 50% |
| Crown Prosecution Service | 100% | | | |
| Legal Aid | 100% | | | |

Source: Office for National Statistics

Determining weights which reflect the relative importance of different dimensions of quality can be difficult. Using the example of courts, is helping reduce reoffending more important than processing cases quickly?

An easier solution is to use an equal weighting for each indicator within the subcomponent. This is the case for all POS subcomponents except prisons.

For prisons, the weightings are 29.2%, 37.5% and 33.3% for recidivism, prison safety and custody escapes respectively. These weights are derived from the Prison Rating System Specification 2014 to 2015 (PDF, 201KB) . In cases where reliable sources on more specific weighting systems than equal splits are available, these should be investigated.

Now that the data are available, limitations of each indicator have been addressed, and weights have been decided on, the calculation of the quality adjusted output measure is possible. One process by which this can be calculated is set out in Table 5. It should be noted that the choice of index number methodology can vary, and this method is just one that could be used.

These data are illustrative only, although have been designed to bear some resemblance to the actual data. As such, the results will not exactly match the estimates in our National Statistic annual release.

Table 5: Producing the non-quality adjusted output percentage changes and quality adjusted output percentage changes for total POS and its subcomponents, from 2015 to 2016

| 2015 to 2016 | Expenditure share of the sub-component | Non-quality adjusted (NQA) output growth | Change in quality adjustment index | Quality-adjusted (QA) output growth |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Prisons | 30% | -0.1% | -9.6% | -9.7% |
| Fire | 23% | 1.0% | | 1.0% |
| Legal Aid | 14% | -8.5% | -5.6% | -14.1% |
| Probation | 10% | 11.2% | -5.6% | 5.6% |
| Crown Prosecution Service | 9% | -5.7% | -5.6% | -11.3% |
| Magistrates | 6% | 0.8% | -0.8% | 0.0% |
| County Courts | 5% | -13.4% | | -13.4% |
| Crown Courts | 4% | -6.9% | -0.8% | -7.7% |
| POS | 100% | -1.2% | -4.8% | -6.0% |

Source: Office for National Statistics

Notes

1. (1) is the subcomponent's share of total POS expenditure for that year in this case, 2015  Back to table

2. (2) is the change in the index of NQA output growth from the previous year to the current, in this case, 2015 to 2016. Back to table

3. (3) is the change in the quality adjustment index for that component from the previous year, with all included quality indicators aggregated together using the weights specified in Table 4. Back to table

4. (4) equals (2) + (3). Back to table

5. NQA and QA output growth for the POS aggregate are weighted averages of the relevant measures for the subcomponents, with weights given in (1). Back to table

6. These data are illustrative only, although have been designed to bear some resemblance to the actual data. As such, the results will not exactly match the estimates in our National Statistic annual release.  Back to table

7. Expenditure shares of each component may not sum to 100% for POS because of rounding.  Back to table

The effects of the weighting via the expenditure shares are clear to see. The change from 2015 to 2016 was negative for all quality adjustments, especially for prisons. This example reflects the findings presented in the latest National Statistic article. Note again that the data in this example are illustrative only, but it is designed to show the same general trends as the real data.

Fire and county courts are not adjusted for quality, so column (3) is left blank for these subcomponents, and (4) is the same as (2).

Note how changes in (3) are the same for subcomponents where the quality adjustments and weights used are the same.

The impact of the actual quality adjustment for POS between 1997 and 2016 can be seen in Figure 1.

**Figure 1: Quality adjusted Public Order and Safety output peaks above non adjusted output in 2009, but experiences a stronger downwards trend since then**

**Non-quality adjusted output and quality adjusted output for Public Order and Safety, UK, 1997 to 2016**

## Figure 1: Quality adjusted Public Order and Safety output peaks above non adjusted output in 2009, but experiences a stronger downwards trend since then

Non-quality adjusted output and quality adjusted output for Public Order and Safety, UK, 1997 to 2016
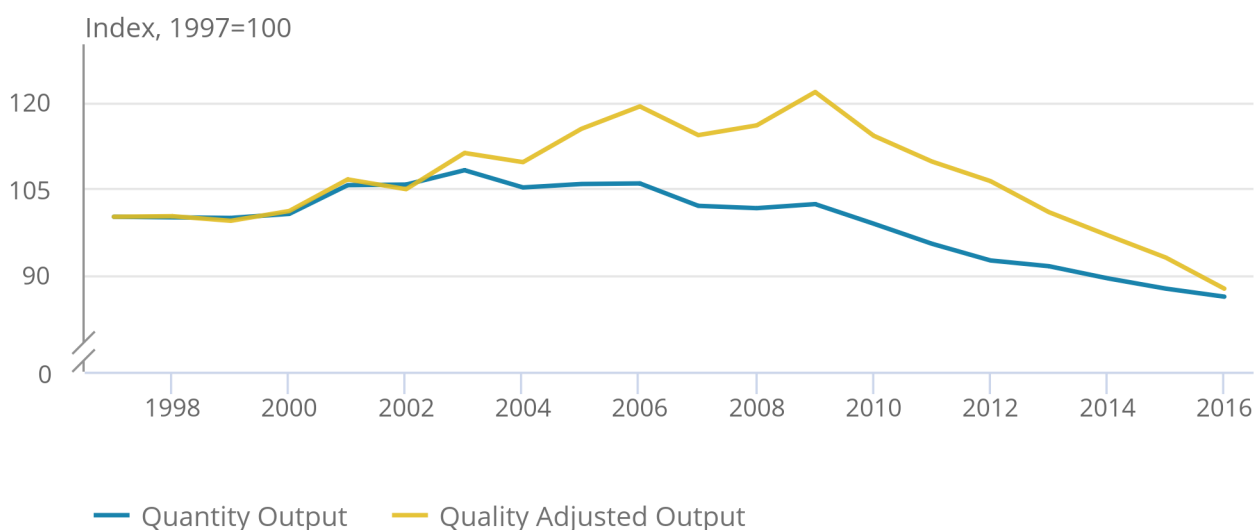


**Source: Office for National Statistics – Public service productivity: total, UK, 2016**

In the annual estimates in our National Statistic release, changes in the quality adjustment and non-quality adjusted (NQA) output are treated symmetrically. That is, the quality adjusted output index can be found by adding together changes in quality and NQA output. For instance, a 2% increase in NQA output and a 1% increase in quality leads to a 3% increase in quality adjusted output. The weighting need not be one-for-one – if improvements in quality were considered more valuable than quantity increases, then this could be reflected in a weighting system that valued quality improvements higher. However, it is not apparent whether, or to what degree, a 1% increase in the number of cases processed or a 1% increase in the timeliness of existing cases is more valuable to society. Given the lack of literature and evidence on this topic, the symmetrical approach is used at the present time.

## Notes for: Designing a quality adjustment – in the context of Public Order and Safety

1. See, for instance, work by other NSIs, such as New Zealand Productivity Commission's Measuring state sector productivity (PDF, 4.4.MB), and academic researchers, such as Growing the Productivity of Government Services.

2. See Chapter 8 of A compendium of research and analysis on the Offender Assessment System (OASys) for more detail on the scale.

# 5 . What do we adjust for quality and how does it affect our statistic?

Quality adjustment measures are implemented in our National Statistic annual estimates, for four service areas: Healthcare, Education, Public Order and Safety, and Adult Social Care. The experimental quarterly series includes no quality adjustment.

The output measure used and the percentage coverage for each service area is shown in Table 6 below.

Table 6: Services areas adjusted for quality in the National Statistic

| Output measure | Service area | Coverage (%) |
| --- | --- | --- |
| Quality adjusted output[1] | Healthcare | 80 |
| | Education | 74 |
| | Public Order and Safety | 75 |
| | Adult Social Care | 41 |
| Quality adjusted "output-equals-inputs"[2] | Adult Social Care | 59 |

Source: Office for National Statistics

Notes

1. Quantity output is adjusted for the quality of the services delivered. If the quality adjustment is positive estimates of output growth will increase. Back to table

2. Calculated in the same way as "output-equals-inputs", an adjustment is subsequently applied to account for changes in the quality of services delivered. While output is still measured indirectly, changes in productivity can be observed as the service quality improves or declines. Back to table

3. The coverage does not sum to 100% for each service area as some output is measured on a non-quality adjusted basis. Back to table

The latest methodology papers for our quality adjustment measures by service area are linked here:

Healthcare: Quality adjustment of Public service health output: current method (PDF, 152KB)
Education: Sources and Methods: Public service productivity estimates: Education
Public Order and Safety: Quality adjustment of Public service public order and safety output: current method
Adult Social Care: Public service productivity: adult social care, sources and methods, 2019 update

The effect of quality adjustments on the public service productivity estimates are notable. In our [latest release](#) (at the time of writing) it was shown that non-quality adjusted public service productivity fell by 3.1% between 1997 and 2016, while quality adjusted productivity rose by 4.0%. This is shown in Figure 2. For a discussion of how quality adjustment affects the total productivity index, see our latest annual release.

**Figure 2: Non-quality adjusted productivity falls over the time series, but quality adjusted productivity increases, with a stronger divergence as time goes on**

**Total public service productivity index, quality adjusted and non-quality adjusted, 1997 to 2016, UK**



Figure 2: Non-quality adjusted productivity falls over the time series, but quality adjusted productivity increases, with a stronger divergence as time goes on

Total public service productivity index, quality adjusted and non-quality adjusted, 1997 to 2016, UK
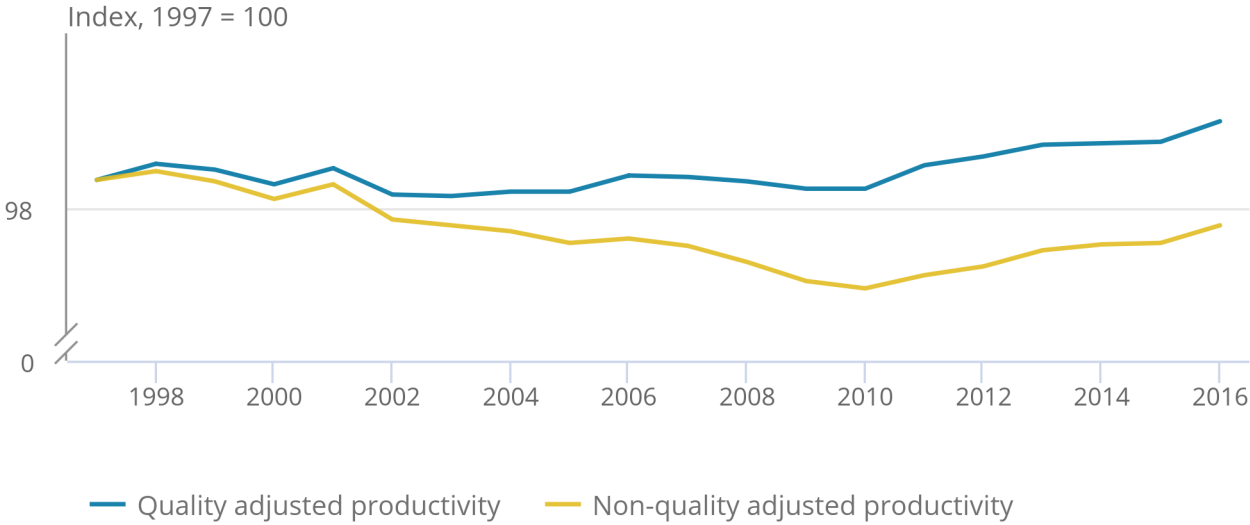
**Source: Office for National Statistics – Public service productivity: total, UK, 2016**

Public service total output has increased at an average rate of 2.6% a year between 1997 and 2016. Quality adjustment has generally improved output growth; however, in 2013 it contributed negative 0.4 percentage points to output growth, leading quality adjusted output to fall. This is shown in Figure 3. The quality adjustment contribution was also negative in 2014, but the non-quality adjusted growth was high enough such that quality adjusted output still grew by 1.7%.

**Figure 3: Quality adjustment tends to improve output growth, with exceptions in 2013 and 2014**

**Contribution to quality adjusted output growth by component, 1998 to 2016, UK**

## Figure 3: Quality adjustment tends to improve output growth, with exceptions in 2013 and 2014

Contribution to quality adjusted output growth by component, 1998 to 2016, UK



Percentage points

● NQA Output   ● Quality Adjustment   — Quality adjusted output
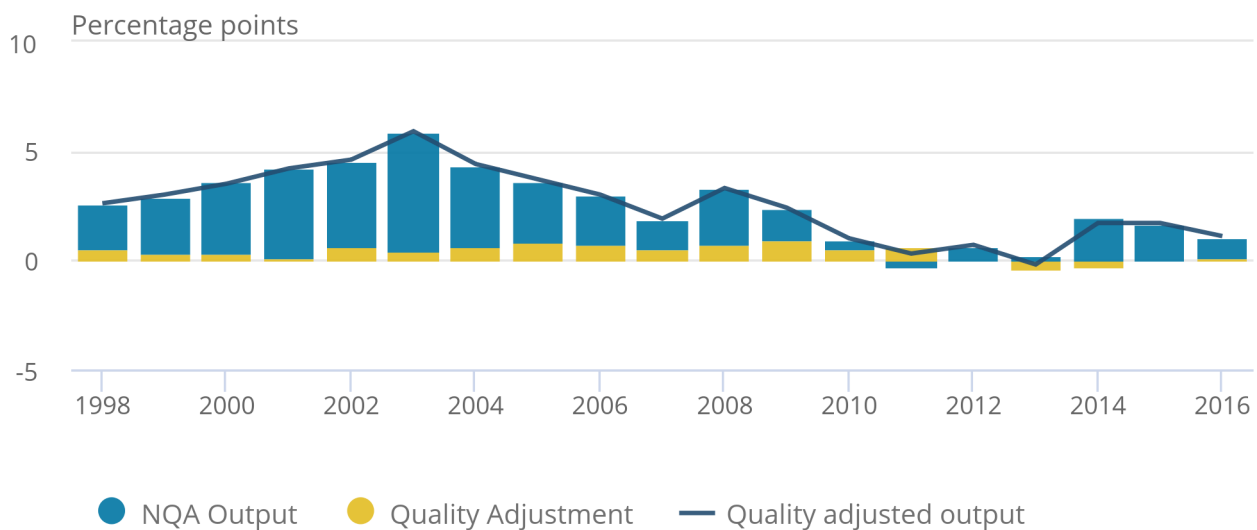
**Source: Office for National Statistics – Public service productivity: total, UK, 2016**

# 6 . A brief history of quality adjustment in the UK

A summary of the history of quality adjustment in the UK is presented in the list below, with some more details to follow:

- Pre-1998: Output was assumed to equal inputs for public service measurement.

- 1998: Direct measurement of output started, following the 1995 guidance issued by the European System of Accounts.

- 2005: The Atkinson Review: Final Report was published, describing nine principles to measure government activity effectively.

- 2006: UKCeMGA (UK Centre for the Measurement of Government Activity) established at the ONS and consultation started.

- 2007: UKCeMGA strategy published.

- 2010: Organisational restructuring and publication of MOPSU (Measuring Outcomes for Public Service Users) final report, the end product of the project originally known as QMF (Quality Measurement Framework).

- 2015: In February the ONS published the first total public service productivity article in its current form.

- 2016: After a Spending Review 2015 bid, ONS re-invested in this area of production, creating new quality adjustments in Public Order and Safety and Adult Social Care, and launching the quarterly series.

Up to 1998, the Office for National Statistics (ONS) measured the output of public services for the National Accounts by assuming the output were equal to inputs. This implies that the level of productivity does not change, and so the growth rate of productivity is always zero.

From 1998, and following European directives, the ONS started to measure public service output by direct methods. The then National Statistician, Len Cook, asked the economist Sir Tony Atkinson to carry out a review of these processes for further development. The Atkinson Review: Final Report (PDF, 1.07MB) was published in January 2005.

Of the nine principles in the Atkinson Report, two were considered particularly important for follow-up work:

- Principle A: the measurement of government non-market output should, as far as possible, follow a procedure parallel to that adopted in National Accounts for market output.

- Principle B: the output of the government sector should in principle be measured in a way that is adjusted for quality, taking account of the attributable incremental contribution of the service to the outcome.

UKCeMGA (UK Centre for the Measurement of Government Activity) was established to carry out this work. It followed up the Atkinson Review by including some quality measures in productivity articles on Education and Healthcare. A consultation was held during 2006 and 2007 to hear from experts and give further legitimacy to the development process. This ended with publication of a strategy in 2007, Measuring Quality as Part of Public Service Output (PDF, 238KB) which set out a conceptual framework for measuring the quality of public services.

Additionally, UKCeMGA secured additional funding from HM Treasury, and worked with various partner organisations[1] , to develop a quality measurement framework for public services. This was a three-year project called the Quality Measurement Framework (QMF). The purpose was to develop effective and easy-to-use methods for measuring the value added of the relevant public services, to be used by government.

The project was set up to research four areas of public service delivery in detail: care homes; knowledge and information services for adult social care; low-level social care interventions, and pre-school education.

In 2009, QMF's name was changed to reflect more accurately what it was delivering – measuring outcomes for public sector users, hence MOPSU. The final report, simply called Measuring Outcomes for Public Sector Users (PDF, 836KB) was published in 2010. It included the Adult Social Care Toolkit (ASCOT) which was later adapted and is still used today to measure the quality of Adult Social Care from a user's perspective.

Since February 2015, the annual estimates of public service productivity have been published in their current format. Since then, continuous development work has led to, among other improvements, two new quality adjustments for Public Order and Safety and Adult Social Care and the design and publication of an experimental quarterly series. The "expansion and review of quality adjustments" is one of the development priorities in the Productivity development plan: 2018 to 2020.

### Notes for: A brief history of quality adjustment in the UK

1. These organisations were: Department of Health; Department for Education and Skills; the National Institute for Economic and Social Research; the Personal Social Services Research Unit at the University of Kent; and National Council for Voluntary Organisations.

# 7 . At the forefront of quality measurement: the international perspective

The Office for National Statistics (ONS) is a world leader in measuring the output of public service provision, and especially in our application of quality adjustments. Other national statistical institutes (NSIs), and research bodies, have also made progress in this area. Relevant findings are described in this section. In the literature, other NSIs variously refer to state, non-market, public sector, or public service productivity – these all refer to marginally different groups of economic activity, but the general objective is the same.

The New Zealand Productivity Commission has produced research on quality adjusting education and healthcare services – see Quality adjusting sector-level data on New Zealand schools (PDF, 804KB) for a good discussion on options for quality adjustment in schools and the greatest challenges surrounding output and outcomes adjustment. Their report Measuring state sector productivity (PDF, 4.4MB) offers some case studies, with clear steps to follow to develop an estimate of productivity (one of the examples is "early childhood education"). Section 7.2 of this report discusses quality in particular, focusing mainly on education and healthcare, and concludes by confirming that there is no easy solution to quality adjustment and that productivity measures are sensitive to the adjustments chosen. However, it notes, you can design appropriate proxies for changes in quality.

The New Zealand Productivity Commission has also produced work on the productivity of courts, Productivity measurement case study: Courts (PDF, 365KB). It suggests using the timeliness of courts alongside the time spent in court per case (sittings or hearings) as adjustments. Finally, Estimating Quality-Adjusted Productivity In Tertiary Education: Methods and Evidence for New Zealand (PDF, 1.3MB) considers the activities of teaching and research in universities, and what some suitable adjusted output measures could be. The latter is not work the ONS has undertaken, as the UK university system is not considered a public service.

The Australian Productivity Commission investigated the productivity of what was called the non-market sector in Supporting Paper No. 2, Non-Market Sector Productivity (PDF, 323KB), one of 16 papers produced for Shifting the Dial: 5 year productivity review in 2017. The need for representative and consistent data was recognised, with efforts to estimate education output constrained in the case of pre-schools, for example. The paper notes that Australia has made good progress in this area in recent years, with direct measures of output the focus. A literature review covers work by the UK, New Zealand, and the United States.

Statistics Denmark engaged with UKCeMGA while it was active, with research such as Quality of Public Health Care and Educational Services (PDF, 204KB) produced. Quality adjustment options for children's and adults' social care, and the commonly discussed health and education, were offered in General Government Output and Productivity 2008-2014 (PDF, 959KB). It was suggested that for social care indicators, the most feasible options of data collection may be through specialist observation and surveys.

Other national statistics bodies appear to produce some form of public service productivity estimate, although quality adjustment specifically is less common. According to Challenges in the Measurement of Public Sector Productivity in OECD Countries (PDF, 306KB), the Netherlands, Portugal and South Africa have made advances in the field. Sweden, according to A Review of the Atkinson Review (PDF, 127KB), has researched it for many years. It is likely that there are others not on this list that have worked on the questions discussed in this guide.

Lastly, measuring the productivity of public service provision in a way that does account for quality is the subject of much academic research. See, for example, Adjusting the Measurement of the Output of the Medical Sector (PDF, 755KB) from the USA's Bureau of Economic Analysis and Public sector productivity: puzzles, conundrums, dilemmas and their solutions (PDF, 970KB). The book Growing the Productivity of Government Services is often referred to by those interested in the field; quality is considered throughout.

The incorporation of quality adjustments into estimates of public service productivity is vital. As shown in Section 5, including quality adjustments can have a significant effect on the total productivity index. This guide serves as an introduction to quality adjustments; there is a wealth of literature available that offers further explanations or unique case studies.

We welcome questions or feedback. Please send your comments to Leah Harris or Josh Martin at productivity@ons.gov.uk.

# 8 . Where can I find more information?

In addition to the resources linked throughout this paper, this section provides the most recent methodology information for each service area that we adjust for quality, and additional papers on quality adjustment for the historical perspective. The current methodology information for each service area is named as such for easy identification. Some of the historical papers are in a general context and some relate to a specific service area. Most were published when UKCeMGA (UK Centre for the Measurement of Government Activity) was active.

## General

Measuring Quality as Part of Public Service Output: Strategy following consultation (PDF, 239KB) presents the results of a UKCeMGA consultation in September 2006 on some key methodological issues, including work on developments on measurement of education and health services.

Following the Atkinson Review: the quality of public sector output (PDF, 122KB), 2007, looks at value weights, a concept raised in the Atkinson Review.

The ONS Productivity Handbook (PDF, 3.2MB), 2007, covers quality adjustment in Chapter 6, page 71.

Adjusting Measures of Public Service Output for Quality of Service (PDF, 175KB) considered some of the issues that were being addressed in order to take the strategy linked first in this list forward, with a theoretical basis. It was published in 2008.

Quality Matters: Update on ONS Methods of Including Measures of Quality as Part of Output and Productivity of Public Services (PDF, 169KB) was published at the same time as the ONS Productivity Handbook. It explains recent developments in work by UKCeMGA to include quality change as part of the measure of public service output and productivity. It sets this work in the context of market sector models on consumer choice and their applications in public service reform.

[The Welfare Implications of Public Goods: Lessons from 10 years of Atkinson in the UK (PDF, 588KB)](#), 2019, offers a history of quality adjustment in the UK, detail on the current methodology, and the challenges around measuring welfare gains in public services and more generally. This paper is published by ONS staff via the Economic Statistics Centre of Excellence Discussion Paper series.

## Healthcare

Current method: [Quality adjustment of Public service health output: current method (PDF, 152KB)](#)

[Public Service Productivity: Healthcare Summary (PDF, 954KB)](#) was released in 2008 to present results of a consultation on how to incorporate Healthcare quality, and a strategy for next work.

[Accounting for the Quality of NHS Output (PDF, 2.2MB)](#) was published by the Centre for Health Economics, University of York, who designed the Healthcare quality adjustment, in 2018. The paper reflects ongoing development work for the adjustment. This paper was used for the list of statistical criteria given in Section 3 and shows that other quality series are being investigated and published.

## Education

Current method: [Sources and Methods: Public service productivity estimates: Education](#)

[Measuring government education output in the national accounts: An overview of several methods developed as part of the Atkinson Review (PDF, 572KB)](#) looks at the National Accounts treatment of education in light of the Atkinson Review, with some ideas for quality adjustment, including Ofsted ratings and cohort progress.

[Public Service Productivity: Education (PDF, 652KB)](#) was a UKCeMGA 2007 article that presented productivity estimates for Education. It also describes the four quality options that UKCeMGA consulted on at the time.

[Quality adjustment for public service education: triangulation (PDF, 141KB)](#), 2012. This paper reviews a range of quality options for Education, including GCSEs after a policy change to the treatment of vocational qualifications in the attainment tables.

[Methods changes in Public Service Productivity Estimates: Education 2013 (PDF, 183KB)](#) was published in 2015 to describe the impact of the Wolf Review on vocational education and the change to the education quality measure.

## Public Order and Safety

Current method: [Quality adjustment of Public service public order and safety output: current method](#)

## Adult Social Care

Current method: [Public service productivity: adult social care, sources and methods, 2019 update](#)

[Public Service Productivity: Adult Social Care (PDF, 184KB)](#), 2007, presents estimates of Adult Social Care productivity and discusses quality throughout, as part of the response to the Atkinson Review. No quality measures were implemented at this point; they were included from 2018, covering 2011 onwards.

## Defence

[Measuring defence (PDF, 196KB)](#) was published in the Economic & Labour Market Review in 2009. It offers alternative measures of inputs and output beyond the National Accounts "output-equals-inputs" convention and also offers ideas for quality adjustment.

## Social Security Administration

[Accounting for Quality Change in Estimates of Social Security Administration Output and Productivity (PDF, 176KB)](#), 2008, considered average clearance times for load and money mispaid for claims as possible quality indicators.

# 9 . Authors

Leah Harris and Josh Martin, Office for National Statistics.