Office for Statistics Regulation

Insight Programme

Approaches to communicating uncertainty in the statistical system

September 2022

The role of the Office for Statistics Regulation

As an independent UK-wide regulator, we are in a unique position to take a broader look at issues of importance to society and to make the case for improved statistics across organisation and Government boundaries. This is supported by our ability to convene, influence and highlight best practice from other sectors.

This review forms part of our Insight programme which, underpinned by the Code of Practice for Statistics, aims to highlight broad themes across the UK's statistical system drawn from our regulatory work.

We want to ensure that statistics provide a robust evidence base for national and local policy development and decision making. We champion the need for statistics to support a much wider range of uses, including, by charities, community groups and individuals. They should allow individuals and organisations to reach informed decisions, answer important questions, make the case for change or hold government to account.

Executive summary

Introduction

Describing, quantifying and visualising uncertainty in statistics can be complex and we have discovered through our work that this can present many challenges for producers of statistics. Ensuring that uncertainty around estimates is conveyed well is an important part of ensuring the appropriate use and interpretation of statistics. At OSR, we are the UK’s official statistics regulator, and we consider uncertainty through all parts of our regulatory work. This report presents our initial findings in this area and is intended to be the first part of an ongoing series of outputs that consider uncertainty in statistics.

Findings

Presenting uncertainty in a meaningful, succinct way that delivers the key messages can be challenging for producers. This is particularly true when it comes to publishing data tables. We found that uncertainty is typically better depicted and described in statistical bulletins and methodological documents than it is in data tables, data dashboards and downloadable datasets.

There has been a rise in the use of infographics and other ways of visualising uncertainty. We saw great examples from some producers of where this has been done well. But in general, more needs to be done in this space as it can be particularly useful when producers show uncertainty in a graphical form.

We also found that there is a wide and increasing range of guidance and advice to help producers think about how to best present uncertainty. OSR will do more to promote and support good practice, and consider what this means for our regulatory work. We will focus on the judgements that we make and the guidance we produce in order to help producers to improve the presentation of uncertainty.

1 Introduction

1.1 Why uncertainty matters

The aim of publishing statistics is to provide insight into the broad range of society's questions. As part of this, these statistics should be useful for their intended purpose. Uncertainty exists in statistics in various forms – for example because of lags in administrative data systems, or through limitations in data collected through sample surveys. An important part of ensuring the appropriate use of statistics, as guided by the Code of Practice for Statistics, is to make it clear that uncertainty exists in the statistics, so that users can avoid drawing inappropriate inferences from the statistics. This requires statisticians to understand and calculate (where possible) measures of uncertainty and to communicate them in a way that can be easily understood by a potentially diverse group of users. Depending on the context, this may be done through appropriate use of language, or simple narrative descriptions of quality, right through to presenting a very detailed quantification of uncertainty. It is especially important to describe uncertainty where there are changes in the quality of the statistics over time, for example as a result of new methods or because of changed data collection approaches, such as the changes we saw during and since the Covid-19 pandemic. This is an area that we will be focussing on further over the coming months. We consider uncertainty as we review statistics as part of our regulatory work but this is the first time we have taken a retrospective look at our work on uncertainty over the past few years to try to draw broader insights.

1.2 What is uncertainty?

Uncertainty has many facets. The Winton Centre for Risk and Evidence Communication at the University of Cambridge distinguishes between direct and indirect uncertainty [1]. Direct uncertainty is where you are expressing your uncertainty about the estimate (or fact), without taking account of any of the caveats that may exist around the way that the data were collected. Indirect uncertainty refers to the uncertainty in terms of the quality of the underlying knowledge that surrounds a claim about a fact, number or hypothesis. This will often be communicated as a list of caveats about the underlying sources of evidence or it can sometimes be summarised into a qualitative or ordered categorical scale such as the GRADE scale for communicating the quality of underlying evidence about the effects of medical interventions. For example, taking the estimate of net migration in the UK in 2021, the direct uncertainty could be summarised by including a range of values within which the true value is expected to lie, assuming a representative sample has been taken. The indirect uncertainty would talk about how those data have been collected, which groups may be missing or are likely to be under-reported and so on.

It can also be helpful to split uncertainty into narrow and broad uncertainty. Narrow uncertainty concerns a specific claim about a defined quantity. It comprises both quantifiable statistical error and (usually unquantifiable) systematic biases due to data limitations. Its expression may take the form of quantified measures where

[1] Van der Bles A.M., van der Linden S., Freeman A.L.J., Mitchell J., Galvao A. B., Zaval L., Spiegelhalter D.J. (2019): 'Communicating uncertainty about facts, numbers and science', Royal Society Open Science, 6(5):181870.

available, or use of words such as “about”. Broad uncertainty relates to the relevance of the number to the wider question of interest. This may take the form of caveats due to ambiguity of terms, a particular metric being a limited measure of the thing of interest, or data simply not being available.

1.3 Communicating uncertainty

Communicating uncertainty is not necessarily an easy task. The relevant aspects of uncertainty need to be presented in a concise and straightforward way that facilitates easy interpretation but doesn’t swamp the statistical messages or paint an overly negative view about the statistics themselves. There are particular challenges in communicating uncertainty in statistical tables (particularly those that are user-defined) and in raw data files. Statistical producers will also likely have varying degrees of understanding of the different aspects of uncertainty, and have to target descriptions at a potentially diverse range of users from different backgrounds and with varying uses of statistics in mind.

At a very basic level, it needs to be very clear when numbers are being presented as estimates. The prominence and visibility of statements on uncertainty is another key consideration. There is less value in having a strong quantified statement about quality, for example, if it is not readily accessible to users.

Descriptions of the uncertainties around estimates can be present in any of a producer's published outputs or online interactive tools. Often, this information can be found in the quality and methods documents. But some indication of uncertainty is also needed in statistical reports, data tables, interactive maps, data dashboards and downloadable datasets to ensure that all users accessing the information understand how to use the statistics appropriately. These challenges are important for producers to address to ensure that their statistics are used correctly – particularly to ensure that any false conclusions are not drawn from the statistics.

Our approach to evaluating the way that uncertainty is described uses the two axes of “what is said” and “where it’s said” to help think about whether information about uncertainty is adequate. Information about uncertainty needs to be both helpful, and presented in such a way that it is accessible to those using the statistics or data. More details on our approach are included in Annex A.

1.4 Aims of the project

In this work we have drawn together what we know about existing guidance and practice across government for communicating uncertainty, along with insights from our own regulatory work. The aim has been to provide a range of examples of good practice to support statistical producers, and to help us improve the way that we regulate. This is only an initial exploration into the topic, and following this analysis, we will work with others to enhance existing guidance where possible and then promote the outcomes to the Government Statistical Service (GSS) and the wider Analysis Function.

We want producers to be equipped to be able to measure and evaluate uncertainty in their statistics. We also want them to have a framework, guidance and good practice examples to be able to use in considering the implications of uncertainty on the use of their statistics, and then to communicate that uncertainty to enhance the use and reduce the potential for misuse of their statistics.

Further details on our approach in gaining insight into our work on uncertainty is covered in Annex B.

The remainder of this report outlines some of the tools and resources that are currently available on uncertainty and then goes on to look at what we discovered from looking across our published work over the last few years. We include a range of case studies.

2 What we found

2.1 Current resources on uncertainty to support the Government Statistical Service

The Government Data Quality Hub (DQHub) provides support for the GSS and wider Analysis Function on the quality of statistics including uncertainty. This is centred around the guidance Communicating quality, uncertainty and change and an associated online course on the Office for National Statistics (ONS) Learning Hub (requires login). The latest version of the guidance was published in late 2018 and the DQHub intends to update this in the near future, taking onboard the insight from this and related work.

One of the goals in the GSS Quality Strategy and the related ONS Statistical Quality Improvement Strategy is that: "We will ensure our data are of sufficient quality and communicate the quality implications to users."

The DQHub manages networks of ONS and GSS Quality Champions, sharing best practice. It ran a GSS sharing webinar on uncertainty (Youtube video) in June 2020, including contributions from analysts in government and from the Winton Centre. One of the presentations showcased the Uncertainty Toolkit for Analysts in Government which sits alongside the Aqua Book and gives a suggested set of guidelines when communicating uncertainty in analysis. The DQHub also provides advice and consultancy, for example collaborating on the guidance from the Race Disparity Audit on Which differences in ethnic group data are real?

In addition to the above, OSR produces guidance that draws out relevant areas from the Code of Practice in relation to Changes in statistical methods and also one that looks at where a change in data quality could mean a change in the National Statistics Status.

Other organisations have also looked at exploring ways of communicating uncertainty. For example:

- In 2020, the Winton Centre wrote about 'The effects of communicating uncertainty on public trust in facts and numbers', which explored and compared whether different ways of communication uncertainty made a difference to the public's trust in the numbers.
- In 2021, FullFact, referencing this earlier work, produced a very useful review on presenting uncertainty including a list of key recommendations.
- ESCOE published research in 2021 concluding that the way that uncertainty information is communicated around productivity measures matters and that by being clear and directly communicating uncertainty was the best approach in setting the public's expectations around future data revisions. ESCOE took this work further and looked at testing different visual representations of uncertainty with the public when comparing international estimates of productivity.

2.2 Uncertainty depicted in statistical bulletins (including visualisations), and methodological and quality documents

2.2.1 Overview

We have reviewed a range of statistics where we have found that producers have described uncertainty well in statistical bulletins. This includes incorporating both technical descriptions of uncertainty along with effective illustrations of where uncertainty lies. We also found that some producers went on to back this up by including links within the bulletin to more-technical methodological articles for those wanting to delve deeper. The best of these technical methods and quality documents would include comprehensive descriptions of the sources of bias and other errors, quantifying them where possible and helpful.

When we have intervened on inappropriate use of statistics, we found that there was sometimes a lack of supporting information covering limitations in the data, which failed to prevent inappropriate use. For example, we found cases where comparisons had been made over time or between countries where the data were not comparable.

Even though quantitative measures such as confidence intervals are sometimes published, they are sometimes presented in isolation without any supporting guidance or context. This can lead to the reported statistic being mistakenly interpreted by some users as the 'true' and only value, rather than an estimate that is likely to fall within a range of possibilities. It can be helpful to use terms such as 'around', 'nearly' or 'about' when presenting figures: these terms can help to inform the reader that the figures are not exact and therefore carry a level of uncertainty. This helps to avoid the risk of misleading users.

We also identified a wide use of rather loose statements concerning uncertainty such as 'figures should be treated with caution'. While these caveats can play an important role in highlighting that limitations exist, in some cases these statements are not specific enough to be helpful and can be overused to the extent that they become lost. This lack of clarity about the limitations can result in either a misuse of a statistic or the opposite, where a user may decide erroneously that it is not safe to use a particular set of statistics at all due to insufficient information about the underlying quality. It is important for statisticians to understand how their statistics are being used and provide accessible information on quality to support their use.

We found that where producers had published information on the quality assurance processes and the risks associated with using administrative data, such as misreporting, users were more aware of how to use and correctly interpret the statistics. Another finding that we often mentioned in our work is that producers should make greater use of the Quality Assurance of Administrative Data (QAAD) framework as a tool to reassure users about the quality of the data sources. Without this, users can sometimes assume that administrative data will always be complete, which may not be true. For example, data will have been collected for a different and specific operational purpose where completeness of some data may have been less important or not possible, or may be incomplete at a point in time because of delays in updating the administrative system.

Uncertainty can also arise from survey data and can be driven by the survey sample size and its bias. For example, uncertainty can come from differential non-response or poor survey design. With administrative source data, bias can also be present in the resulting statistics. This might be through changes to the variables collected in the administrative dataset or decisions around the inclusion of records with missing data. A decision to omit these records from the analysis could introduce a bias in the statistics.[2]

Our analysis identified a tendency to report information about statistical biases insufficiently clearly in statistical bulletins. Examples include not reporting on the level of non-response, coverage or whether self-reporting or reporting on behalf of someone else. These types of biases can be present in statistics based both on survey and administrative data sources.

There has also been an increase in the linking of datasets to create richer data sources. This is a positive development, but the linking of these datasets may introduce biases – for example, if linkage rates for a certain age group or ethnic group are higher than in other groups, then that group could be more represented in the analysis and could skew the results.

In more recent years, particularly in light of the pandemic, more use of modelling techniques has been made as producers try to address issues around reduced sample sizes or changes in mode of collection. Sometimes, the underlying assumptions in the modelling have not been made clear enough or described in an accessible way to allow users to fully understand the implications on the statistics.

When producers report on statistics that are subject to revision, the details are sometimes included only in the methodology reports and it is not always clear in the bulletins that figures are ‘reported’ incidents or events and therefore may be subject to a degree of regular revision. Related to this is the trade-off between what is often referred to as ‘rough and ready’ estimates that are produced at speed and usually with a higher degree of uncertainty and those statistics that are less timely but perhaps more certain. Communicating this distinction and the scale of any revisions is also important for users to know how they should be using the statistics.

The publication of infographics alongside statistical releases has been increasing in recent years and where they are done well, these can provide an easy and quick way to grasp the key messages including where these statistics can best be of use. In the best practice examples we found producers using infographics or introductory bullet points to set out clearly what the data can and can’t be used for. This includes, for example, highlighting where uncertainty associated with different breakdowns may affect the ability to use the data in a certain way especially where comparisons are being made. Specifics in this area included highlighting the increased uncertainty associated with disaggregated data for lower-level geographies. It is particularly useful when producers show this uncertainty in a graphical form.

Fan charts, such as those used by the Bank of England, are a helpful way to communicate inherent uncertainty. Other ways of demonstrating the range of possible values of a projection such as presenting variant projections are also helpful.

The Winton Centre also found that any graphical representation (error bars, fan charts, diffusion plots) of uncertainty around time series maintained trust in both the

[2] https://spiral.imperial.ac.uk/bitstream/10044/1/61527/2/Statistical%20challenges%20of%20administrative%20and%20transaction%20data%20FINAL%20version.pdf

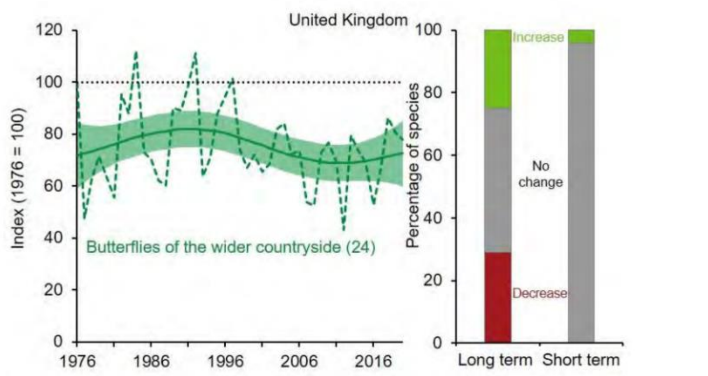
numbers and the producer of the statistics. We recognise that producers have taken steps to use newer approaches to data visualisation to try to overcome these challenges.

2.2.2 Case studies

Department for Environment, Food and Rural Affairs (DEFRA) England Biodiversity Indicators bulletin (Figure C6ib)

DEFRA, within its most recent publication looking at biodiversity, has shown clearly through the bulletin both in the narrative and illustratively that there is uncertainty in the estimates that they present.

Figure C6ib Trends for butterflies of the wider countryside in the UK, 1976 to 2020.



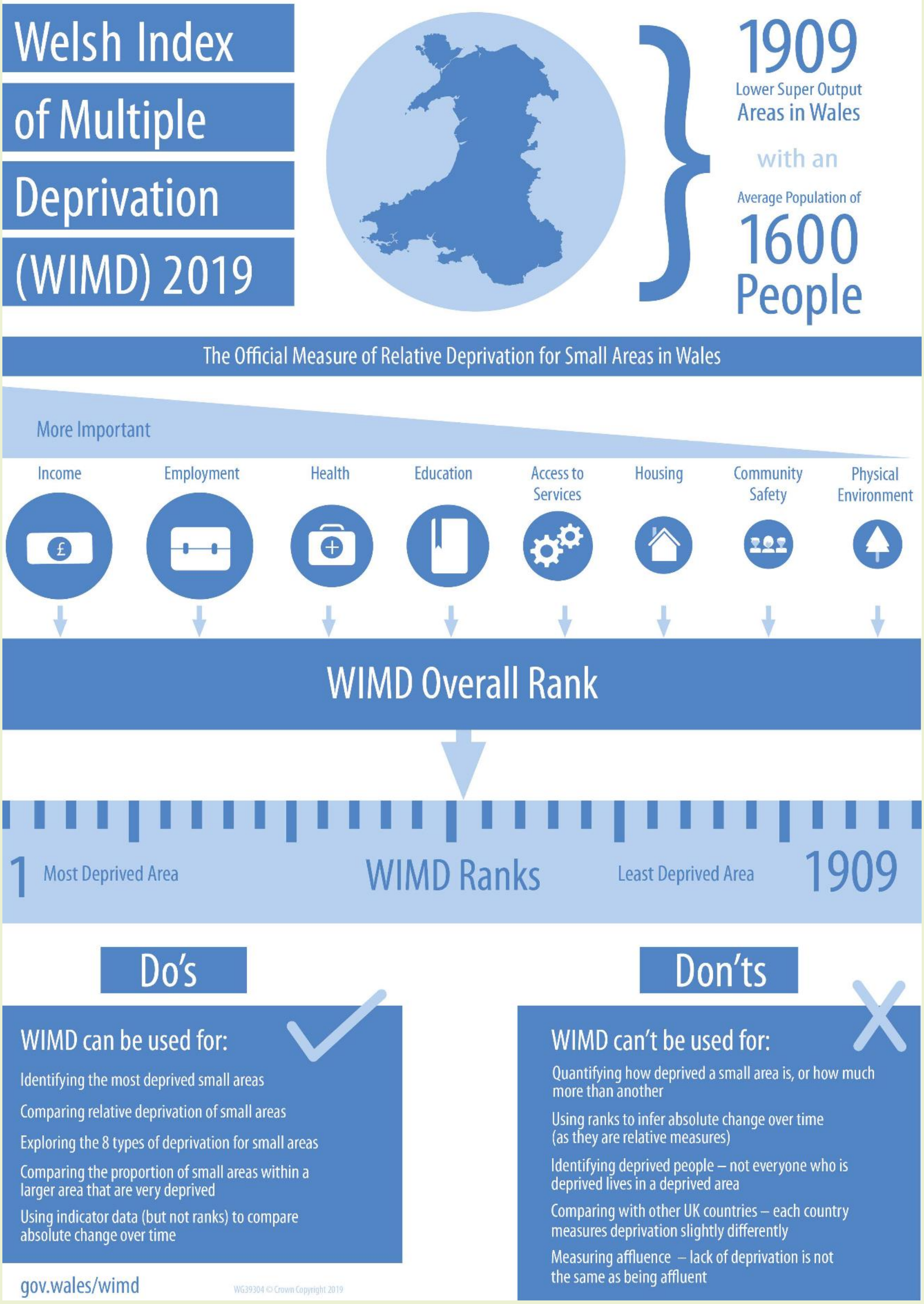
This image shows trends for butterflies of the wider countryside in the UK between 1976 and 2020. The line graph shows the unsmoothed trend (dashed line) and the smoothed trend following modelling (solid line) together with its 95% confidence interval (shaded). The figure in brackets shows the number of species included in the index. The bar chart also included shows the percentage of species within the indicator that have shown a statistically significant and those with no change.

Office for National Statistics: Covid ad-hoc analysis

In November 2020, we published our findings on some casework we received about an ONS ad-hoc analysis of the number of school workers, key workers and other professions in England who had Covid-19. We concluded that ONS had not intentionally presented the analysis in a misleading way but that there were some changes that could have been made to the analysis and the accompanying text to support those reading the bulletin in better understanding the results at that time. We thought that ONS could have done more to explain the uncertainty around these estimates particularly where they concerned education staff categories, and the ongoing implications of this uncertainty. In terms of particular phraseology and use of language, we highlighted that it could have been clearer that ‘no evidence of a difference’, as stated in the bulletin, is not the same as ‘evidence of no difference’. Following publication on this topic (February 2021) ONS addressed these points. In 2020, outputs from the survey were still in their early stages of development and we welcome the continual review and improvement that has been made to statistics produced from the Covid Infection Survey.

Welsh Government: Welsh Index of Multiple Deprivation

The Welsh Index of Multiple Deprivation (WIMD) statistics are widely used by central and local government and community organisations to target services. In addition to a really valuable set of outputs, during this review, we were particularly impressed with Welsh Government's WIMD infographic depicting a complex product but clearly indicating how these statistics can be used and interpreted correctly by all users. The infographic clearly shows what WIMD can and can't be used for in a very visual way.



The Welsh Index of Multiple Deprivation (WIMD) 2019 is the official measure of deprivation for small areas in Wales. It identifies areas with the highest concentrations of several different types of deprivation.

There are 1909 Lower Super Output Areas (or small areas) in Wales, each with an average population of 1600 people.

There are eight types of deprivation (domains) included in WIMD 2019 and these are weighted in order of importance. Income is the most important determinant of deprivation, followed by Employment, Health, Education, Access to Services, Housing, Community Safety, and Physical Environment.

Each domain has at least one underlying indicator. There are 47 underlying indicators used in WIMD 2019. Ranks for the eight separate domains are created by combining relevant indicators within each domain. The WIMD overall ranks are created by combining the domain ranks.

WIMD ranks all small areas in Wales from 1 (most deprived) to 1,909 (least deprived).

Do’s – WIMD can be used for:

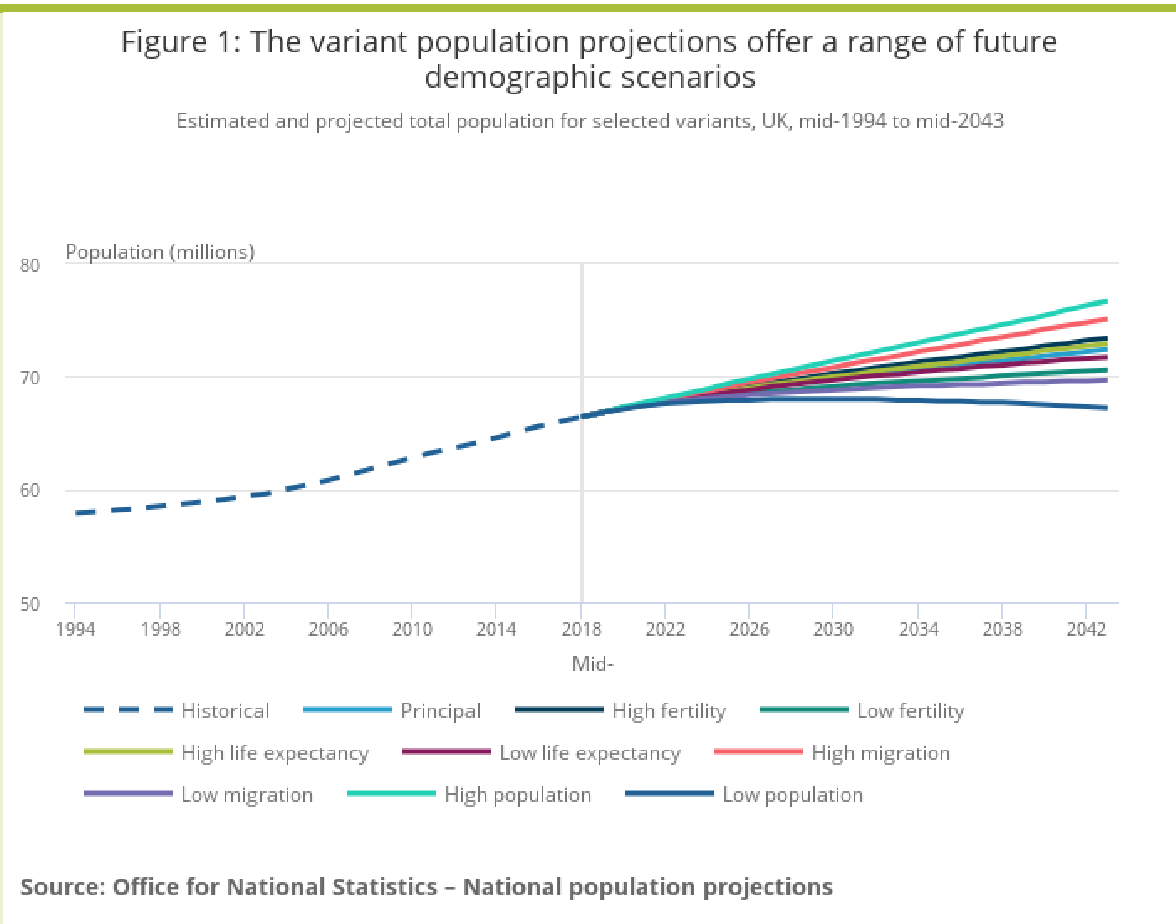
- Identifying the most deprived small areas
- Comparing relative deprivation of small areas
- Exploring the 8 types of deprivation for small areas
- Comparing the proportion of small areas within a larger area that are very deprived
- Using indicator data (but not ranks) to compare absolute change over time

Don’ts – WIMD can’t be used for:

- Quantifying how deprived a small area is, or how much more than another
- Using ranks to infer absolute change over time (as they are relative measures)
- Identifying deprived people – not everyone who is deprived lives in a deprived area
- Comparing with other UK countries – each country measures deprivation slightly differently
- Measuring affluence – lack of deprivation is not the same as being affluent

Office for National Statistics: Population Projections

ONS has developed a helpful way, as shown below, to present the effects of different underlying assumptions that could be fed into a projections model. This is a useful way to show that there is not only one projection but many alternative scenarios and also serves as a reminder that the sources and assumptions going into a model impact the output from that model.



This image shows different population projections (known as variants) and how these compare with the principal projection. These variants will show a different trajectory over time depending on the underlying assumptions that are fed into the model. For the SVG image of this chart (accessible) and original data please see the Figure 1 on the ONS National population projections, variant projections: 2018-based.

2.3 Uncertainty depicted in data tables

2.3.1 Overview

Although statistical bulletins are still accessed and used by a wide range of users, statistics are increasingly being presented and consumed through tables, including detailed data tables charts, infographics. Opportunities are increasing for accessing individual and low-level data either publicly or through restricted-access channels, and these are particularly useful for further detailed analysis. There are specific challenges where users can compile their own data extracts through the use of table builders.

However, although many users are accessing data through tables and table builders, we found, in general, that the presentation of uncertainty in data tables was a particular weakness, even where uncertainty in the corresponding statistical bulletin had been presented well. Many data tables did not clarify that the statistics are estimates or reference uncertainty at all. In the case of projections, data tables were not always clear on what a central estimate was and that this does not necessarily equate to the most accurate point within the confidence band. Although this was a fairly recurring finding in our work, we recognise that it can be more challenging for producers to flag uncertainty while maintaining accessibility of the statistics. They are faced with increasing calls to limit the size of their outputs and publishing estimates that are clear and easy to find without the inclusion of lots of footnotes and caveats is difficult.

Where statistics producers are doing this well, it is through presenting high-level information on uncertainty and appropriate use at the forefront which includes in data tables.

Another issue can be that despite the inclusion of confidence intervals (CIs) in the data tables, users can be left unsure of how they should interpret the estimates and CIs if some high-level guidance is not included alongside.

2.3.2 Case studies

Welsh Government: Motoring Offences Statistics

These data tables are published by Welsh Government and show figures for various driving offences. For example, breath test statistics are used to measure the effectiveness of drink–drive campaigns by police forces.

Although the main statistical release published alongside the data tables is clear that the figures are estimates, there is no mention of uncertainty in the data tables (nor in the statistical quality tab associated with the data tables).

Office for National Statistics: Annual Survey of Hours and Earnings (ASHE)

The Annual Survey of Hours and Earnings (ASHE) tables contain estimates of earnings for employees by sex and full-time or part-time status. An example of one of these tables for gross median pay for part-time employees is shown below. ASHE is based on a 1% sample of jobs taken from HM Revenue and Customs' Pay As You Earn (PAYE) records.

The colour coding within the tables indicates the quality of each estimate based on the coefficient of variation (CV) of that estimate. The CV is the ratio of the standard error of an estimate to the estimate itself and is expressed as a percentage. The smaller the coefficient of variation the greater the accuracy of the estimate. Estimates with a CV greater than 20% are suppressed from publication on quality grounds in addition to those that present a disclosure risk.

This example of indicating differences in quality in a table by the use of colour coding is a good idea, it is both simple and effective and well explained. The downside is that, as with all attempts to demarcate with cut-offs, it can be quite a crude measure and doesn't take into account the uses to which the data might be put, and the level of accuracy that a user might need. There are also questions about ensuring that such an approach would meet website accessibility criteria.

Percentiles									
10	20	25	30	40	60	70	75	80	90
3,744	6,356	7,280	8,180	9,625	12,969	15,138	16,609	18,509	24,975
6,894	8,784	8,844	9,516	11,599	15,155	19,318	22,275	25,786	x
6,856	8,784	8,796	9,492	11,119	15,399	19,875	23,163	26,795	x
x	8,968	9,606	10,246	12,000	14,807	x	x	x	x
6,394	10,592	12,500	14,666	18,266	24,341	27,276	29,246	31,770	40,398
8,635	9,600	12,184	12,500	15,755	24,392	28,526	31,663	34,524	x
6,087	11,266	13,782	15,749	19,017	24,271	27,265	28,991	31,050	37,716
4,982	9,730	12,206	14,627	17,930	24,543	26,864	28,893	31,832	42,226
8,000	10,025	12,471	13,377	17,514	23,910	27,789	29,583	32,234	x
4,471	8,101	9,210	10,162	11,963	15,927	18,476	19,812	21,511	26,299
7,604	10,045	10,520	11,439	12,750	15,918	18,172	19,796	21,573	x
3,557	7,204	9,187	10,643	12,243	15,467	17,153	18,814	20,082	x
x	7,633	8,373	9,798	11,962	20,091	25,032	26,026	x	x
1,018	2,011	2,613	3,445	5,727	9,199	10,781	11,844	12,963	x
7,773	9,792	10,701	11,584	13,320	17,731	20,031	21,661	23,816	29,297

Key	Statistical robustness
CV <= 5%	Estimates are considered precise
CV > 5% and <= 10%	Estimates are considered reasonably precise
CV > 10% and <= 20%	Estimates are considered acceptable
x = CV > 20%	Estimates are considered unreliable for practical purposes
.. = disclosive	
: = not applicable	
- = nil or negligible	

Department for Education: Key stage 2 statistics

In the 2019 Key Stage 2 statistics published by Department for Education, we found that although the tables presented included confidence intervals for the progress scores, there was no context provided and no explanation given about the confidence intervals or how to interpret them. It is positive that the confidence intervals are shown but the inclusion of more context could help users to better interpret what these intervals mean especially where they might be making comparisons.

Office for National Statistics: Low Carbon and Renewable Energy Economy statistics

ONS publishes Low Carbon and Renewable Energy Economy statistics that include a bulletin and data tables. Within each data table, the level of uncertainty associated with the figures is presented as a coefficient of variation (CV), a measure of the relative variability of the data. This gives an indication of some aspects of the uncertainty in the estimates and can aid in the interpretation of the statistics as the estimates and the associated uncertainty are presented together in the data tables where many users will access the raw figures. However, although it is really positive to see an example of where a measure of uncertainty is reported together with the estimates in the data tables, its interpretation is unlikely to be accessible to users who may not be so technical and it avoids reporting on other biases that could be present.

3 Conclusions and next steps

3.1 Conclusions

It is clear that showing uncertainty in estimates, for example through data visualisation, is essential in improving the interpretation of statistics and in bringing clarity to users about what the statistics can and cannot be used for. At the same time, however, we recognise that this is often not always a straightforward task.

We found that uncertainty presentation was best-developed in statistical bulletins. This often comes in use of words like “estimate”, rounded numbers and warnings of caution. In some cases, the warnings to users could be more helpful if they were more specific.

However, one of our discoveries was the relatively low level of reporting uncertainty in data tables. This is a clear gap as many users of statistics will only reference the data tables and extract the data to use for their own analysis. If the level of uncertainty is not evident then further misunderstanding could result. But we also recognise here that the task isn’t easy and we would encourage producers to adopt approaches such as the data shading illustrated earlier as a promising way to making improvements in this area.

Our regulatory work follows suit – our focus to date has been more on bulletins and less on data tables, and there is clearly more that we can do to challenge and support the statistical system in presenting uncertainty across the whole range of statistical and data outputs.

We also found that there is a good deal of guidance already existing to help statistics producers understand and present uncertainty. There is also a range of organisations – the Winton Centre, DQHub and Full Fact to name three – involved in enhancing understanding and developing presentation of uncertainty.

3.2 Next steps for statistics producers

With support from us and those at the centre of the GSS, we encourage Heads of Profession for Statistics to review whether uncertainty is being assessed appropriately in their data sources, and to review how this is presented in all statistical outputs.

As part of this, sharing good practice across the GSS on what has worked well in terms of communicating uncertainty will bring benefits right across the statistical system. One of the key routes to share this good practice will be through the data quality champions network and we encourage the network to support this endeavour. In terms of feeding into work on uncertainty across the GSS and beyond, finding good examples of where uncertainty has been presented well or described well by producers is important. This can serve both as a way of highlighting good work and also showing less experienced statistics producers ways of presenting uncertainty in their statistics that they may well not have thought of.

3.3 Next steps for us as regulators

We will continue to review the communication of uncertainty in our regulatory projects. We already have a good range of experience and effective guidance to help review uncertainty presented in statistical bulletins and methodology documents. We will continue to use this, and enhance as needed.

We will generate new guidance for ourselves to help us evaluate the presentation of uncertainty in charts, infographics, data tables and other “non-bulletin” presentation of statistics. We will use the examples identified so far to help us do this. We will also reinforce the benefits of using the QAAD framework to understand uncertainty associated with administrative data.

We will continue to collect examples that show both good communication of uncertainty and also that might require further work. Through this we can improve on the judgements that we make and the guidance that we produce and start to focus in on more specific areas where improvement is either needed or good work can be showcased. We will work with Heads of Profession for Statistics, and GSS networks (such as the Quality Champions) to help spread and reinforce good practice across the GSS.

We will also continue to work with other partners, particularly DQHub, to strengthen and enhance the current guidance to cover:

1. The presentation of uncertainty in data tables
2. Best practice on using data visualisation to communicate uncertainty
3. Uncertainty in administrative data

We will also work with the Analysis Function to develop guidance around how best to present uncertainty in a way that meets accessibility guidelines. Initially, this would require some scoping work to understand the technical barriers that exist with approved chart tools. This work would also benefit from engagement with the Winton Centre, to understand to what extend its work can be applied within the current accessibility guidelines.

Annex A

OSR Approach to Compliance Checks on Uncertainty

OSR uses the two axes of “what is said” and “where it’s said” to help think about whether information about uncertainty is adequate taking into account what kind of decisions or further analysis the statistics might be used in, and by whom (including their level of expertise). In many cases there’ll be different types of people making different types of decisions, which we bear in mind. In some cases we may need to make some assumptions about the types of decisions made, the types of people or organisations making them, and their level of understanding of the statistics. We find it helpful to think specifically about the potential for misuse (for example drawing a conclusion or making a decision that may not be borne out if we had perfect data) in any given context.

What is said

1. **quantitative description** of uncertainty. This would often be expressed as confidence intervals, margins of error or sampling errors, modelling errors, statistical significance levels. Or could be presented visually such as by using error bars or fan charts or some other shading to represent the likelihood and magnitude of uncertainty.
2. description of the likelihood and potential magnitude of **revisions**. This could either be anticipating (numerically or descriptively) future revisions, or provide an analysis of past revisions.
3. **qualitative description** of possible sources of uncertainty. This category would include basic statements reminding the reader that the data are from a sample survey, and could include descriptions of definitional issues, coverage issues, response biases, any issues relating to time lags etc. It may also include details about the likelihood, direction and magnitude of any possible biases.
4. **use of words** like “estimate”, “approximately”, “about” or “around” that express some uncertainty in estimates. Words like “probably”, “possibly” or “may”, particularly when comparing possible changes over time or differences between categories might be used. Use of rounded numbers also helps avoid spurious accuracy (but the motivation here might be for such as confidentiality protection).
5. **no mention of uncertainty**. This would be where the statistics and data are presented as if they were absolute facts “The unemployment rate was 4.5%”, “GDP grew by 0.2%” etc

Where it’s said

A. **visible and prominent** (you’d have to try hard to miss it) – uncertainty is either presented up-front at the start of the document (or each section, for example) or is presented alongside the numbers or charts
B. **there, but could be more obvious**. This is the kind of scenario where there’s a footnote that might be missed on a casual read, or where there is information in the notes at the end of a document

C. **invisible or so hidden that you'd have to try hard to find it**. This basically covers scenarios where there is no information about uncertainty at all, or relies on a detailed read of a document, following links that aren't very visible or well-named etc.

Annex B

What we did

The aim of this project is to build our understanding of how OSR can support statistics producers to improve the communication of uncertainty within their outputs. It is not our role to prescribe how statistics producers should do this but to highlight good practice where we find it and common areas of challenge. The analysis underpinning this insight project was structured in three parts:

- What guidance on communicating uncertainty is available to statistical producers
- What recommendations has OSR made around uncertainty through casework interventions and regulatory reviews
- What approach are statistical producers taking to describing and presenting uncertainty within their statistics

The Winton Centre for Risk and Evidence Communication Centre has done a lot of great work around communicating uncertainty including its seminal paper on the topic. We reviewed the work that the Winton Centre did and used an adapted form of the Winton Centre's uncertainty framework: our approach is summarised in Annex A. We applied this framework when carrying out our compliance checks as part of our work for this review. This gave us a pool of evidence about how uncertainty of official statistics is communicated. As part of our ongoing regulatory work, we are now formally recording the way that uncertainty of official statistics is communicated.

As part of this current project, we reviewed the existing guidance that focusses on Communicating Quality, Uncertainty and Change guidance for the GSS. This guidance aims to support producers to be able to provide assurance to users on these three areas, explaining complex concepts whilst being clear and transparent about professional judgements. We also explored what other guidance is available in the Government Analysis Function more widely and found the Uncertainty Toolkit website, which is an analyst's guide to dealing with uncertainty which forms part of the Aqua Book resources, to be a useful resource.

In order to analyse what recommendations OSR has previously made concerning uncertainty, we used a combination of web-scraping and database interrogation techniques to search for specific terms related to uncertainty. The web-scraping was carried out on OSR's website to review all published correspondence containing these search terms. We then searched our internal casework database for these terms within emails from correspondents and our responses to them. Due to the potential overlap in published correspondence and responses to casework in the database, a manual review of flagged cases was carried out to remove any duplications. Another limitation of our analysis is the selection of search terms as they may have missed relevant correspondence and publications.

For the final area of analysis, we organised an OSR-wide session to review a range of statistical publications for their presentation of uncertainty. Members of OSR were asked to review the statistics, including the bulletin, tables, methodology documents and any other related outputs, for the presentation of uncertainty against the Winton

Centre’s scoring criteria highlighted in the introduction. 80 sets of National Statistics were chosen at random for review. As we did not conduct a formal sampling approach in our selection of statistics to review, our findings are only indicative and may not fully reflect a producer or the statistical system’s approach to communicating uncertainty.

Using the outcomes from these three parts of analysis, the project team brought the findings together to identify common themes in the approach taken to communicating uncertainty within the statistical system. The project team then discussed the findings with DQHub and Winton Centre to help form recommendations for next steps.