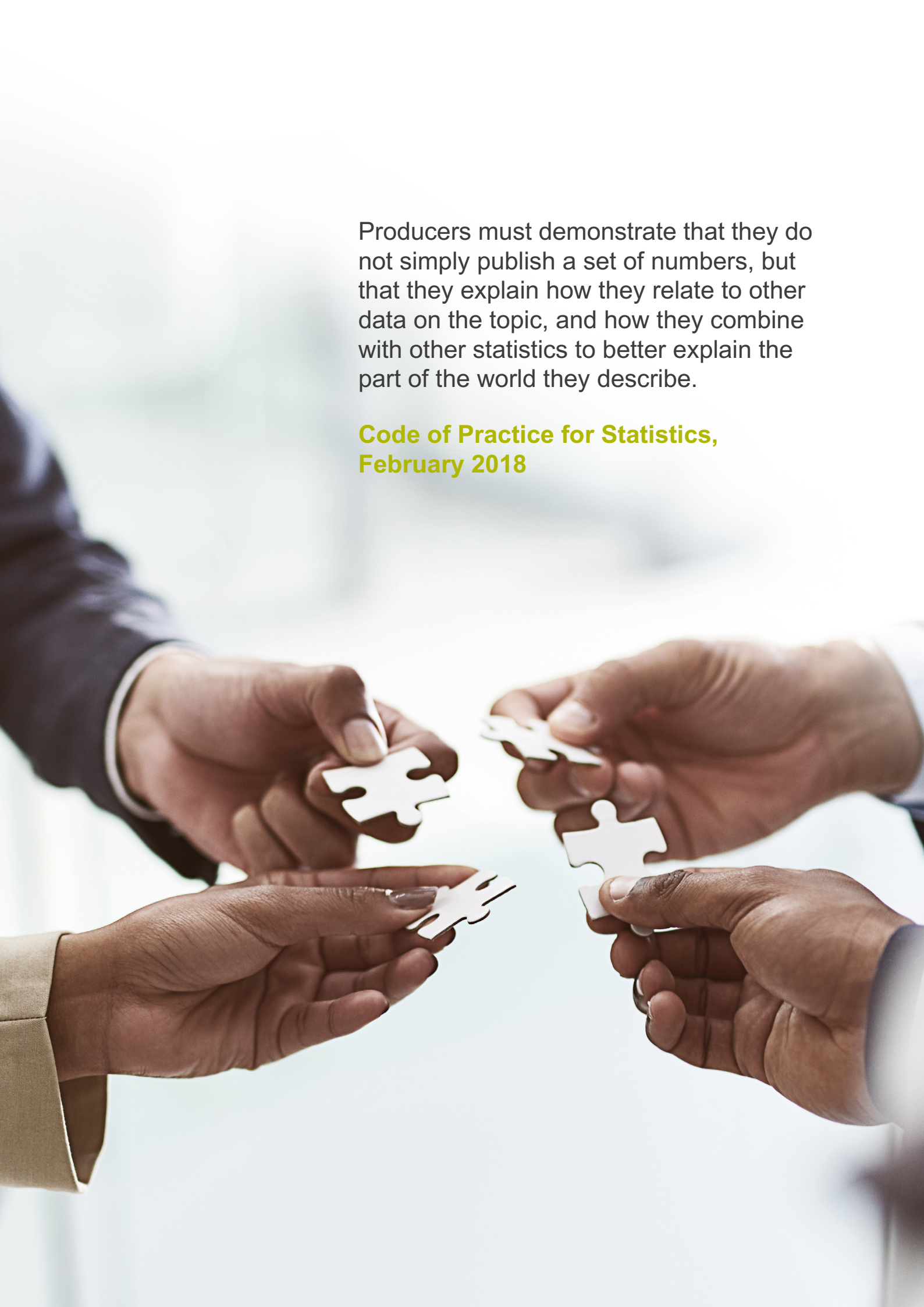
Office for Statistics Regulation

OSR Insight: Coherence

July 2019

Producers must demonstrate that they do not simply publish a set of numbers, but that they explain how they relate to other data on the topic, and how they combine with other statistics to better explain the part of the world they describe.

Code of Practice for Statistics, February 2018



1. Introduction

The Office for Statistics Regulation (OSR) plans to produce a series of OSR Insight reports to share lessons from our regulatory work with a wider audience in an accessible way. This report is the first in that series.

1.1. OSR Insight on coherence

This report focuses on coherence and outlines a framework to support analysts when developing statistics and analysis. OSR has recently published guidance which defines coherence as reflecting the degree of similarity between related statistics and the fuller insight achieved by drawing them together. In relation to coherence, the Code of Practice for Statistics states that:

> *Producers must demonstrate that they do not simply publish a set of numbers, but that they explain how they relate to other data on the topic, and how they combine with other statistics to better explain the part of the world they describe.*

This statement guides our expectations on the commitment producers of statistics should make to demonstrating coherence.

In many cases, analysts – including statisticians across the Government Statistical Service (GSS) – deliver high quality, insightful analysis which answers key questions and is set in context. This report highlights some of these positive examples and sets out a framework (see box 1.1) to support consideration of coherence as statistics and analyses are developed and disseminated.

Box 1.1: Achieving Coherence

Identify the question

Understand the questions users and potential users want answered.

Use appropriate data

Collaborate across boundaries to gain access to data, align definitions, bring sources together and understand differences.

Tell a story

Create a narrative that draws together information to provide clarity and insight.

We hope this approach will complement other work in this area, such as the recently published GSS Harmonisation Strategy – which sets out ambitions around comparability and coherence – and the work of the GSS Strategy Delivery Division. By working together we can highlight the value of coherent outputs and support greater coherence in published statistics and analysis.

2. Lessons from OSR regulatory work

The Code of Practice for Statistics emphasises the importance of coherence and highlights its application across all three pillars: Trustworthiness; Quality; and Value (see box 2.1). It is one of three cross-cutting themes, alongside collaboration and transparency.

Box 2.1: Coherence in the Code of Practice for Statistics

Relevant Code Principles

Trustworthiness

T2 Independent decision making and leadership
T4 Transparent processes and management

Quality

Q1 Suitable data sources
Q2 Sound methods
Q3 Assured quality

Value

V1 Relevance to users
V3 Clarity and insight
V4 Innovation and improvement
V5 Efficiency and proportionality

Guided by the Code, we have considered the coherence of official statistics outputs across our regulatory work. Coherence has been a focus of some recent systemic reviews – which offer an opportunity to look across statistics producers and products – and is an important consideration in assessments and compliance checks.

Our systemic review programme covers broad themes, such as skills, housing, and health and social care, often looking at how the public value of groups of statistics can be maximised and highlighting suggested improvements. We have also included relevant requirements in recent assessments, for example, in our assessments of cancer survival statistics and the Department for Digital, Culture, Media and Sport (DCMS) Sectors Economic Estimates. Coherence has also been referred to in compliance checks and published casework letters, for example in our interventions on accident and emergency waiting time statistics. While we have identified different areas for improvement in each instance our work to date has highlighted several common themes:

- Coordination and high-level strategy can support producers to address the challenge of producing coherent outputs across multiple organisations. We have found that statistics are often produced in silos with a focus on a single output.
- We have seen greater value from statistics where efforts have been made to improve harmonisation, for example, across countries within the UK or at different geographic levels.
- Coherence does not just apply when multiple datasets are being considered but is also relevant to statistics produced from a single source. These statistics need to be considered in the context of the world they describe and other available statistics.
- Users of statistics are better supported when an overarching narrative is presented, for example, pulling together multiple sources to provide an overall story and putting the statistics in context.
- Users of statistics tell us they welcome better information on how different but related statistics fit together, for example, statistics based on different measures or relating to different countries.

3. Identify the question

To be of value statistics must address questions to which people – including government, charities, businesses or individuals – want answers. It is not enough to simply report statistics because the numbers are available and have always been produced. Published outputs should answer questions in a way that supports understanding and adds to the public debate. There are a number of considerations that can support this approach.

3.1. Engage

To identify relevant questions, it is essential to engage with users and potential users of statistics and to monitor public debate (for example, see box 3.1). This will help identify what the statistics should measure and any gaps in available outputs.

Box 3.1 National Records of Scotland Population Projections

National Records of Scotland (NRS) presents its Projected Population of Scotland estimates as responses to a series of topic based questions, such as "Why is Scotland's population projected to increase?" and "How does Scotland compare with the rest of the UK?". These questions are identified by understanding user questions and the key topics of public debate in Scotland.

The population projections are based on the most recent population estimates together with assumptions about future levels of fertility, mortality and migration. A range of data sources, expert advice and user input feeds into the production of the population projections. By bringing together all of this information, NRS is able to tell a coherent story about how Scotland's population is projected to change.

3.2. Think beyond individual products

It is often not possible to answer the question of interest to users through a single source of data. There may be multiple sources of data answering the same question in different ways or using different definitions. Answering the question may require a judgment based on a range of sources.

Box 3.2: Crime Statistics

Questions like “how much crime is the adult population experiencing?” or “is there now more violent crime?” sound straightforward but answering them is not easy.

Different sources of data are more appropriate for the different crime types, so to answer these broader question requires a judgment based on multiple sources. The Office for National Statistics (ONS) and Home Office have been addressing this by focusing on the key questions of interest and presenting answers based on a range of sources.

The OSR blog having a better public debate about crime provides more detail on some of the challenges and how they have been addressed.

The example in box 3.2 demonstrates how multiple sources have been used to address very specific questions. Producers of statistics may also need to consider multiple sources of data to answer a broader question that could not be covered through a single output (see box 3.3 for example). These questions will not always fit neatly within organisational boundaries, so it is important to think of issues in new ways and to collaborate across organisations.

Box 3.3: Living Longer

The Living Longer series produced by ONS considers a variety of pertinent questions that policy-makers, commentators, think-tanks, academics, charities and the public may have about the ageing UK population, irrespective of traditional departmental responsibilities. The series has been designed as a resource for these users to draw on when considering how the UK should adapt to the opportunities and challenges associated with a changing population.

The first article in the series draws on a wide range of government and non-government sources to answer questions of relevance to the economy, public services, wider society and the individual. Subsequent articles have looked in more detail at the interplay between health, caring and working in later life. The most recent article in the series examines the relationship between population ageing, economic dependency and international migration in the UK. Future planned articles will look at older households and housing and perceptions of crime among older people.

4. Use appropriate data

Analysing appropriate data is essential to answering questions and providing a coherent narrative. It is supported by collaboration to gain understanding of the data, align definitions and bring sources together. Analysts need to work across professional, organisational and international boundaries, with coherence often supported by joint analysis plans.

4.1. Coordinate analysis plans

It will often take more than one organisation to answer the questions people care about. To achieve a coherent message these organisations need to work together on a coordinated approach to analysis. This could include jointly commissioning outputs that can give a fuller picture of an individual, for example Sport England's Active Lives Children and Young People Survey, or working together to coordinate analysis plans, for example in housing (see box 4.1).

Box 4.1: Housing Statistics

Within the housing and planning statistics landscape there are at least 11 government departments, devolved administrations and public bodies across the UK that produce relevant official statistics. In response to the OSR housing review a cross-government housing statistics group has been established to provide leadership and improve collaboration across producer organisations.

The group has developed and published a joint work programme and is working towards delivering greater coherence through outputs like the UK Private Rented Sector report published in January 2018. It brings together existing UK data sources on the private rented sector to assess comparability, coherence and data limitations. Similar cross-UK articles on homelessness and affordable housing are due to be published in Autumn 2019.

4.2. Align definitions

There are many aspects to aligning definitions. For example, harmonising survey questions, comparability across levels of aggregation (geographic or other), consistency over time, and using methods and definitions that are consistent with international best practice.

Harmonisation of definitions should be aimed for, but where harmonisation is not possible it is important to work across organisations to understand differences or consider how to produce alternative measures which allow comparison. Some of the challenges in achieving harmonisation are set out in the homelessness definition work undertaken by the GSS Harmonisation team. One common issue is the different policy contexts encountered in different countries across the UK. Box 4.2 shows how this has been addressed for accident and emergency waiting times.

Box 4.2: Accident and Emergency Waiting Time Statistics

Due to differing policies and targets in different nations, A&E waiting time statistics have not been available on a comparable basis across the four countries of the UK. In September 2018, NHS Digital published official statistics about attendance and waiting times in A&E in England on a basis comparable to estimates for the other three countries of the UK.

The release includes detail of the data sources used for each country and provides helpful links to the official statistics each one publishes. The publication demonstrates collaboration by statisticians representing all four countries of the UK. More information is available in our report on NHS performance measures.

4.3. Bring sources together

Coherence often requires drawing on a range of sources of data and information. This could be record-level or aggregate data to enable analysis, provide context or support validation. The data could be produced by government or the private sector and could include traditional or newer approaches, for example, surveys or web-scraped data.

The OSR Joining Up Data report highlights some powerful examples of data linkage being used to provide insights and drive policy change. It highlights the importance of coherence and consistency and outlines a range of areas where more could and should be done to use data linkage to help answer important questions. The new GSS web pages on data linking, developed in response to the Joining Up Data report, offer examples of good practice as well as practical resources to support analysts and departments.

Box 4.3 summarises work done by the Department for Business, Energy and Industrial Strategy (BEIS) to access, understand and link household-level data.

Box 4.3: National Energy Efficiency Data-Framework

BEIS has worked with a range of organisations to gain access to data to produce the National Energy Efficiency Data-Framework (NEED). Most of the data in this framework are not owned by BEIS, but BEIS has a unique role in being able to draw together the property-level information to answer important questions about energy efficiency and energy consumption. The outputs have been used to produce analyses and disseminate information to support decision making by government, individuals and businesses, including through tools like the non-gas map and table creator.

The work of the statisticians is done in collaboration with departmental scientists, which allows statistical analysis and scientific studies to be considered together to support policy on savings from installation of energy efficiency measures. BEIS has also produced a record-level dataset based on NEED to allow wider analysis of the data.

4.4. Understand the data

Producers of statistics must understand the data they are working with. This includes understanding definitions, quality and how the data reflect what they are attempting to measure. This is essential in identifying the most appropriate data source and ensuring users are aware of any limitations. It is only by understanding the data that producers of statistics and analysis can understand which questions they can accurately answer and explain what the data represent.

The Quality Assurance of Administrative Data (QAAD) standards set out OSR's expectations about understanding administrative data and offers guidance which can be relevant to producers in other scenarios. It highlights the importance of understanding the context of data collection and undertaking comparisons with other data sources.

5. Tell a story

The approach to dissemination of statistics and analysis is an important aspect of coherence. Outputs should draw together information to provide clarity and insight.

Creating a narrative which guides users through key points can be achieved in a range of ways, including traditional reports, linked webpages, visualisations or interactive tools. The complexity of the information and the way users engage with it will determine the best approach. Whatever the output there are some common features that can support coherence.

5.1. Put data in context

An effective statistical product should take the reader through the information highlighting key points and putting data in context. This should support the reader to understand the implications of the findings and how they can inform decisions. To do this effectively, statisticians need to be aware of related analysis produced by other organisations (inside and outside government) and key issues in the wider public debate.

One approach to providing context is pulling together relevant statistics in a combined publication which supports interpretation of data in a more meaningful way. This could draw together a range of data from one organisation into a single output, such as the Welsh NHS activity and performance summary, or pull together data and evidence from a range of sources such as the Department for Environment, Food and Rural Affairs' (Defra) Future Farming and Environment Evidence Compendium (see box 5.1).

Box 5.1: Future Farming and Environment Evidence Compendium

Defra's Future Farming and Environment Evidence Compendium brings together official statistics, economic analysis, social research, operational research, and scientific analysis to tell a joined-up story about the current state of agriculture in the UK in an accessible way.

Another aspect is to explain consistency and comparability with other related statistics. For example, if two sources of data appear to answer the same question then it is important to explain what the differences are and when it is appropriate to use each of the sources. An example where this has been done clearly – with the data source being put in context – is by Welsh Government via a blog explaining measurement of Welsh language use and Welsh speakers.

5.2. Coordinate dissemination

Working with others to coordinate how statistics are published can support understanding of a topic and increase use. This could be working together to produce a report that answers questions of interest, or pulling together multiple data sources in one place to support their use. It often requires thinking outside organisational boundaries. See boxes 5.2 and 5.3 for examples.

Box 5.2: Race Disparity Unit

The Race Disparity Unit ethnicity facts and figures website – which has voluntarily applied the Code of Practice for Statistics, including through a 'users-first' approach – provides data about the experiences of people from different ethnic backgrounds. It gathers data collected by government and others with more sectors still to come, on topics as diverse as education, employment, health, housing and criminal justice in one place, making data available to everyone, including the public, academics, specialist practitioners, businesses and charities.

Box 5.3: Health and Social Care Statistics

The English Health and Care Statistics Steering Group (EHSSG) was set up in response to OSR's systemic review. The group aims to improve comparability and accessibility of health and social care statistics through closer collaboration across producers.

The appetite for change is strong, and the formation and close working of theme groups have enabled statisticians to learn from each other, achieving better engagement with external users and reducing overlap in some published outputs, such as smoking statistics. It has led to the publication of new information on the GSS website, including a comparison of key measures and a positively received interactive web-based tool that allows users to search for all the outputs on a health or social care topic.

Learning from EHSSG is being shared with the four-nations Health and Care group to promulgate good practice across the UK.

6. Conclusion

Achieving coherence is not always straightforward, but the value in doing so is significant. Statistics and analysis can be used more widely with greater confidence, leading to better decision making and more informed public debate.

It is therefore essential that analysts continue to work towards producing more coherent outputs. This report outlines an approach to considering coherence and highlights examples of good practice. OSR will continue to champion coherence and work with others to enhance coherence in statistics and analysis produced by government.

Office for Statistics Regulation

Statistics that serve the public good

OSR Vision, 2019

If you have any feedback on this report or suggestions for future OSR Insight topics please email **regulation@statistics.gov.uk**

www.statisticsauthority.gov.uk/osr
