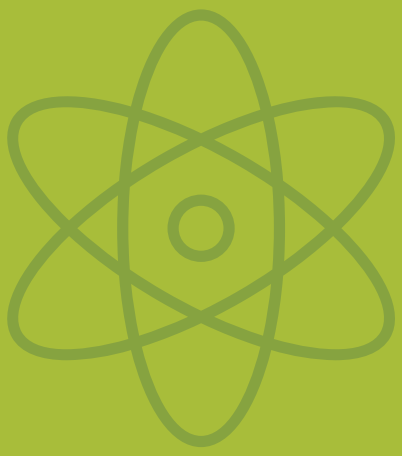
Office for Statistics Regulation

Office for Statistics Regulation
Research Programme



The Public Good of Statistics: What we know so far

December 2020

Contents

Executive Summary

In 2019, the Office for Statistics Regulation (OSR) created a research programme dedicated to developing a better understanding of how statistics can serve the public good. To succeed in this aim it is critical to understand what is already known, and what is not known, about the public good of statistics. This literature review will help us to achieve that aim. It will also help to guide future research plans.

In this review, we begin by considering how the public good can be defined, measured, and conceptualised. We then consider the legislative framework which defines the public good of statistics in law. Following this, we review empirical research which is relevant to the question of whether statistics are currently serving the public good.

We identified two important themes which require further research to develop a better understanding of whose needs are currently served by statistics: trust in statistics and statistics producers, and statistical literacy. Evidence suggests that these two issues may be instrumental in ensuring that statistics can serve the widest range of users.

The review also considers the economic value of statistics. This approach highlighted two important factors: the need for measurements which can quantify the value of statistics and the need for more timely statistics on economic measures.

Next, the review considers the social value of the public good of statistics by discussing the impact of data gaps and evaluating whether the approach taken by the BBC to provide a valuable service to the public can offer insights and possible comparisons to OSR's approach to the public good. Finally, we end the review by summarising our findings and outlining the future of OSR's research programme.

The Office for Statistics Regulation

Statistics shape our understanding of the world and help us make important decisions about our lives. The increasing availability of data has only supported this further but having more statistics does not necessarily mean having better statistics, or more useful statistics. The Office for Statistics Regulation (OSR) has a mission dedicated to upholding the trustworthiness, quality, and value of statistics, and to protect the role of statistics in public debate. To achieve this, OSR independently regulates statistics: OSR assesses whether they comply with the Code of Practice for Statistics, promotes adoption of the Code's principles more widely and intervenes when issues with statistics have been identified. However, the scope of OSR's mission goes beyond this.

In order to fulfil OSR's vision of statistics serving the public good, OSR intends to develop a better understanding of what this means. To do this, OSR established a research programme in 2019 which has the purpose of gathering evidence to understand how well statistics meet the needs of users. OSR will not achieve this vision alone, and it is not alone in wanting statistics to serve the public good; the research programme will involve collaboration with a wide range of stakeholders who may have different approaches to the subject. By engaging with this work, the research programme aims to transform the way that producers think about statistics, so that they are better informed about how their statistics can best serve a very wide range of users.

Aims and Objectives of the Review

The research programme has defined research questions which will be addressed through the course of its programme of work. The first of these research questions is 'whose needs are served by statistics?'. The **objective** of this review is to take the first step towards answering this important question. This review will gather and compare evidence to create insights which will influence the development of the research programme.

In this review, the **aim** is to identify current approaches to the concept of the public good of statistics, and to understand what is known, and what is not known, about this concept. This is not intended to be a systematic review; this literature review intends to create a better understanding of how public good can be conceptualised and the different approaches which have been taken to apply this term. This review aims to create a map of the landscape of the public good to clarify this term and to assess how the research programme can best contribute towards this. Firstly, key definitions will be described.

Key Definitions

The Public Good

A key purpose of the research programme is to understand precisely what is meant by statistics serving the public good. There is an established definition in economic literature for **a public good**; this commonly refers to a resource which is non-excludable, meaning everyone can use it, and non-rivalrous, meaning one person

using it does not take it away from anyone else (Nicholson, 1978). Statistics which are free to use, such as official statistics, meet this definition. However, OSR has an ambition for statistics to be elevated further than meeting this basic principle so the benefits associated with statistics can be maximised and felt by the widest range of users. OSR refers to this as **the public good.**

It is also important to note that there are numerous phrases used in literature which refer to the same concept, or a very similar concept. For instance, the public good may also be referred to as common good, public benefit, societal benefit, public interest, or social value. Waind (2020) also notes that, while previous studies have given some indications as to what the public perceives as the public good, this concept does not seem to have a widely understood and accepted definition amongst the public. They suggest that individuals may have different ideas about what this means exactly, identifying that there is a clear need to establish how the public view the public good.

Statistics

Statistics refers to "the understanding of information contained in data" (Grenander & Miller, 1994) or "the technology of extracting meaning from data and of handling uncertainty" (Hand, 2008, p.2). The remit of OSR's regulatory work is official statistics, which means statistics produced by the government's departments and agencies. The term 'National Statistics' however can only be conferred by OSR and this refers to statistics which are assessed as being fully compliant with the Code of Practice for Statistics. National Statistics are therefore a subset of official statistics. In this review, the term statistics will be applied broadly to refer to any statistics, rather than just official or National Statistics.

Statistics serving the public good

For statistics to serve the public good, we must turn what is envisaged into best practice. Florence Nightingale wrote that "The main end of statistics should not be to inform the government as to how many men have died, but to enable immediate steps to be taken to prevent the extension of disease and mortality" (Nightingale, 1858, as cited in McDonald, 2012, p. 328). In this sense, statistics are used practically in the performance of a service, or to contribute towards public delivery, in a way which does not discriminate against individuals but instead seeks to maximise the benefits of statistical research. Statistics serving the public good is now defined in legislation and refers specifically to informing the public about social and economic matters and assisting in the development and evaluation of public policy. To OSR, statistics serving the public good also means;

- Statistics upholding the principles of trustworthiness, quality and value
- Statistics which meet the needs of the widest range of users
- Statistics which are communicated effectively

Approaches to the Public Good of Statistics

In this review, four approaches to defining and evaluating the public good of statistics have been identified, which will be discussed in turn:

- **Legislative approach**, which creates a legal framework to underpin the importance of the public good being served by statistics, and to enable mechanisms for this to happen
- **Empirical research**, which uses research to understand the views of the public on statistics and related issues
- **Economic value**, which outlines the importance of timely and accurate statistics, as well as being able to measure the monetary value of statistics
- **Social value**, which looks at how society benefits from comprehensive statistics

Within these approaches evidence will be discussed which relates to statistics and the public good, and gives an overview of the broad landscape in which this term is applied.

Legislative Approach

The public good of statistics is a concept which began to gain prominence in the UK in the 1990s. Prior to this time, statistics were regarded as a resource primarily intended to serve governmental decision making. However, attitudes to statistics began to change over time until they became viewed as a resource which should be available for everyone to use. Due to the increasing importance of the availability of high-quality official statistics, legislation was developed to protect the integrity of the statistical system and to embed statistical independence. The first piece of legislation to do this was the Statistics and Registration Service Act (2007).

Statistics and Registration Service Act (2007)

In 2007, the Statistics and Registration Service Act became law. This law created an important legacy as it reinforced statistics as a protected public resource. It also led to the creation of the UK Statistics Authority, which is an independent body at arm's length from the government that oversees the Office for National Statistics (ONS, the largest producer of official statistics in the UK) and OSR, which is its regulatory arm.

The UK Statistics Authority is tasked with the objective of "Promoting and safeguarding the production and publication of official statistics that serve the public good". The legislation then goes on to define what this means: "Serving the public good includes in particular — (a) informing the public about social and economic matters, and (b) assisting in the development and evaluation of public policy". This cemented the definition of official statistics serving the public good in legislation, an important milestone in protecting and upholding the role of statistics for the public.

Digital Economy Act (2017)

The concept of statistics serving the public good was further supported by a piece of legislation in 2017 known as the Digital Economy Act. This law created mechanisms to facilitate data sharing and linking which means researchers and statisticians can submit applications to the National Statistician's Data Ethics Advisory Committee (NSDEC) and the UK Statistics Authority Research Accreditation Panel (RAP) to make use of official statistics and other public sector administrative data. These mechanisms therefore help to increase the insights gained from data collection by making data available to qualified people. Furthermore, to be successful in this process, applicants need to demonstrate how their research will serve the public good, signifying the importance of this concept in legislation.

Empirical Research

Empirical research concerning statistics often does not explicitly ask participants whether statistics serve the public good. The following surveys provide insights into perceptions of statistics (and statistics producers), engagement with statistics, and whether statistics are valued by the public.

Public Confidence in Official Statistics (PCOS, NatCen, 2019)

The Public Confidence in Official Statistics report (PCOS) is based on the British Social Attitudes Survey which aims to measure how trust in official statistics has changed over time and how much awareness the public have of ONS. Nearly 2000 people in England, Scotland, and Wales participated in the survey.

When asked about their awareness of ONS, most participants (69%) had heard of it with varying degrees of awareness; for example, 17% of participants said they knew ONS 'well', 30% of participants said they knew ONS 'somewhat', 22% had only ever heard the name, and 30% of participants said they had never heard of ONS (1% of participants responded that they did not know). The results also demonstrated a high level of trust in ONS. When asked if they trusted ONS, 24% of participants said they did not know but, excluding these participants, 88% responded that they trusted ONS 'a great deal' or 'tend to trust it', leaving 12% who said that they tend to distrust ONS.

To explore attitudes towards official statistics, participants were asked whether they trusted ONS statistics. Similar to finding that a majority of participants trusted ONS, a majority of participants (85%) also said they trusted statistics produced by ONS (excluding 19% of participants who answered that they did not know). Furthermore, considering the reasons behind this trust: 28% of participants said they trust ONS statistics because 'ONS does not have a vested interest in the results', however 23% of participants said they distrust ONS statistics because 'the Government has a vested interest in results', which may suggest that some participants are unaware that ONS is a non-ministerial government department.

In terms of the value placed on ONS statistics, a high majority of participants (93%) agreed that statistics are 'important for understanding our country', indicating that they place a high value on ONS statistics (excluding 18% of participants who were not sure if they were important or not). However, the findings showed that, even

though official statistics are valued by the majority of participants, 76% of participants stated they have never used or referenced official statistics. This raises an important question about the extent to which official statistics 'should' be used, or directly accessed, in order for us to be confident that the public good is being served.

Examining this point further, the demographic information about participants demonstrated a pattern which showed that the participants in the highest socio-economic group and the participants in the highest level of education group were more likely to use statistics than participants in the other groups. These participants were also more likely to trust ONS and their statistics compared to other participants. Perhaps this suggests that not being exposed to statistics, or opportunities to engage with statistics, may mean that statistics and statistics producers are more likely to be viewed with suspicion.

One important point to explore further is to improve our understanding of the direction of the relationship between statistics and trust. What the findings presented do not tell us is whether engaging with statistics through education or occupation increases one's trust in statistics, or whether that trust is required to motivate an individual to engage with statistics. The importance of trust was the focus of the next study which will be discussed, which looks further at the perceived trustworthiness of statistics producers.

Veracity Index (Ipsos MORI, 2020)

The Veracity Index is the longest-running survey on trust in professions in the UK (it has been administered since 1983) and it has the aim of measuring how much the British public trust different professions. Whilst this survey does not explicitly ask about statistics, it measures the levels of trust in the professions who produce, publish, and communicate statistics and can therefore provide some valuable insights. For this survey, 957 British participants were asked, 'For each profession, would you tell me if you generally trust them to tell the truth, or not?' in relation to 30 professions.

The results indicated that 60% of participants said they trusted civil servants (who produce official statistics) to tell the truth, which represents an increase of 35% since 1983. This means that civil servants were ranked 13 out of 30 professions in terms of their trustworthiness, putting them just above the middle of the rankings. However, the demographic information of participants demonstrates education-based differences, such that those who have a degree were more likely to state that they think civil servants will tell the truth (69%), compared with those who do not have a degree (43%). This is reminiscent of the findings of the PCOS which showed that participants in the higher education group were more likely to say they trust ONS. It is unclear why this education-based difference exists but it may be worth future research investigating whether this is a barrier to statistics serving the public good.

Looking at the Veracity Index more broadly, other statistics producers and communicators do feature on the list: above civil servants are professors (83%, 6/30) and scientists (82%, 7/30) and below civil servants are journalists (23%, 27/30) and pollsters (53%, 17/30). Considering that these professions may all be responsible for the production and/or communication of statistics, it may be useful to consider why there are differences in perceived trustworthiness.

This is a pattern which has also been found in other studies. Research carried out by the Winton Centre (2020) measured changes in levels of trust in professionals communicating with the British public during the COVID-19 pandemic. This online survey demonstrated that trust in scientists and science professionals was consistent from March to October, but trust in politicians, the current government, and civil servants (or policy officials) was relatively lower compared to science professionals and it decreased over the same period. This demonstrates a similar pattern to the Veracity Index as science professionals appear to be trusted more than policy professionals or the government. This has been further evidenced by an Ipsos MORI (2020) poll which found that participants believed that politicians should pay more attention to experts when making decisions compared to paying attention to civil servants, the findings from experiments, or their constituents.

In light of all of this evidence, it may be useful to consider why these professions differ so much in terms of trustworthiness. Understanding more about the qualitative differences between these professions may provide insights into this pattern. As previously discussed, this may be a critical point for research if a lack of trust in the communicator is a barrier to statistics serving the public good. Additionally, developing a better understanding of the impact that the trustworthiness of the communicator has on the information being delivered may be a very valuable question to consider in future research. In the next survey, the public good of statistics is considered in relation to trust and the value placed on statistics.

The Value of Official Statistics (United Nations Economic Commission for Europe [UNECE], 2018)

The Conference of European Statisticians (CES) created a global task force of experts to investigate how valued statistics are. The Task Force commissioned national statistical offices (NSOs) to complete a survey which asked users if statistics are of value to them, how they could be of more value, and how the value of statistics could be better communicated. The findings from this survey demonstrated that the perceived trustworthiness of statistics related to how valued the statistics were, highlighting the lack of value in statistics which are not perceived to be trustworthy. The findings again reinforce the importance of trustworthy statistics when considering whether statistics are serving the public good.

The Task Force also shone light on the difficulties that non-expert users can face when navigating official statistics, with results suggesting that non-expert users increasingly want to read ‘the story’ of the statistics, rather than numbers without context. This may be because evidence also emerged from the NSOs suggesting a lack of statistical literacy in users, as shown through the misuse, or lack of use, and misunderstanding of official statistics.

Statistical literacy has been defined by Wallman (1993) as “the ability to understand and critically evaluate statistical results that permeate our daily lives — coupled with the ability to appreciate the contributions that statistical thinking can make in public and private, professional and personal decisions” (p.1). If statistical literacy is lacking amongst non-expert users, as suggested by evidence from the NSOs, statistics producers, or intermediaries such as journalists, must carefully consider the most effective way to communicate statistical information. Murray and Gal (2002) suggest that producers of statistics, amongst others, have an important responsibility to

ensure that the public can comprehend statistical publications. The same could potentially be said about the communication of economic information, as can be seen in the next study.

Public Understanding of Economics and Economic Statistics (Runge & Hudson, 2020)

Research from the Economic Statistics Centre of Excellence (ESCoE) (Runge & Hudson, 2020) used 13 focus groups (N = 130) and an online survey (N = 1665) to investigate public perceptions and understanding of economics and economic statistics. The findings demonstrated that some participants in the focus groups expressed a sense of distrust and apathy towards the economy. The authors suggest this may stem from economic information being communicated in, what is perceived as being, an inaccessible way. If the information is communicated using unnecessary jargon, this may be alienating to non-expert users, therefore the authors suggest further research is needed to understand how these messages can be communicated more clearly.

The findings from the survey also suggested that there were qualitative differences in how knowledgable and confident participants were in understanding economic concepts. Participants who were male, older, had higher socioeconomic status, and a higher education level were typically more knowledgeable and confident in this area. These participants were also more interested in economics, which may imply they have more motivation than others to understand the economy. However, some participants in the focus groups stated that they found economics important and interesting and regretted not having a better understanding of it. These participants suggested that perhaps economic concepts should be taught more widely in schools, which may be an interesting avenue to consider in terms of addressing economic literacy. In the next survey, education and confidence with numbers are considered in the context of understanding causes of death.

Perils of Perception (Ipsos MORI, 2020)

The Perils of Perception survey aims to measure how accurate participants are in estimating the causes of death in their country. In this survey, 16,000 participants in 32 countries were asked, ‘Out of every 100 deaths [in your country] – about how many are a result of [cause of death]?’, with causes of death referring to cancer, terrorism or conflict, and transportation injuries, for example. The findings demonstrated that there was high inaccuracy across all countries and some consistent patterns emerged, for instance deaths by cancer and cardiovascular diseases tended to be underestimated but deaths from terrorism or conflict, or transportation injuries, tended to be overestimated.

The authors considered multiple explanations for these findings: for example, whether participants had been affected by the causes of death listed, or whether they see the causes of death frequently in the news. These factors had some bearing on guesses at a country level, but the authors suggest that it is much more complex to understand why participants were inaccurate at the individual level. One explanation the authors investigated was whether being ‘comfortable working with numbers’ related to accuracy; their analysis demonstrated a weak positive

correlation wherein participants with higher confidence in using numbers more likely to be accurate.

On a related point, throughout the history of the Perils of Perception survey, Duffy (2018) states that there is a consistent clear positive correlation between education and accuracy. This pattern was established at the country level by ranking countries according to their results for the OECD's Programme for International Student Assessment (PISA). At the individual level, Duffy (2018) states there is also a clear positive correlation between education and accuracy, wherein those with the highest levels of education tend to be more accurate. Duffy stresses that there are limits to what can be inferred from this correlation and therefore further research is needed to understand this pattern.

Still, it is important to acknowledge the significance of this pattern; without a clear understanding of the risks to life, the public may not develop a full appreciation of how best to maintain their health. For instance, living under the impression that there is a high chance your death could be caused by terrorism, or injuries incurred in transportation accidents, may generate a skewed view of society and undue anxiety about daily activities. Similarly, having the impression that there is a low chance of death by cancer or cardiovascular disease may mean that health advice is disregarded and unhealthy lifestyles are maintained. Whilst it is important to avoid overextrapolation, it is clear there is a link between being comfortable working with numbers and having a proportional awareness of risk which does require investigation.

Susceptibility to Misinformation About COVID-19 Around the World (Roozenbeek, et al., 2020)

Further related evidence has recently emerged which supports this pattern. Roozenbeek and colleagues (2020) investigated factors related to susceptibility to misinformation about COVID-19. Participants were asked to complete numeracy tests (which tested their actual numerical ability rather than their confidence in using numbers) in addition to being presented with statements about COVID-19, some of which were common examples of misinformation, and asked to rate how reliable the statements were. Participants were also asked further questions about their compliance with health advice, what sources they seek information from, and their levels of trust in scientists.

The findings suggested three main factors related to susceptibility to misinformation about COVID-19. Firstly, trust in scientists was negatively associated with susceptibility to misinformation, wherein participants who stated they had higher trust in scientists were less likely to find the examples of misinformation to be reliable. Secondly, participants who sought information from the World Health Organisation were less likely to be susceptible to misinformation, compared to participants who sought information from social media. Thirdly, participants who performed better on the numeracy tasks were also less susceptible to misinformation compared to participants who did not perform as well. This evidence further underlines the point that numeracy may be critical to understanding how important information can be effectively communicated to non-expert users.

Early Conclusions Based on Empirical Research

The importance of **trustworthiness** is clear in the findings presented; statistics must be seen as being trustworthy in order to be considered valuable by participants. However, the evidence also showed considerable differences in terms of how trustworthy statistics communicators are perceived as being, particularly in the Veracity Index. Future research could investigate the impact that this variation in trustworthiness has on the message being delivered to contribute towards our understanding of how information can be effectively communicated.

The findings also demonstrated that there are education-based differences in who finds statistics trustworthy and who finds statistics producers trustworthy. These education-based differences may indicate that only some participants feel that statistics are serving their needs, but it is clear that further empirical research is warranted to understand why these differences exist in terms of trust. This is especially important as work by Roozenbeek and colleagues (2020) showed that participants who trust scientists more are less likely to be susceptible to misinformation, and they are also more likely to be compliant with health advice. Understanding how this trust can be established and maintained is a very important question for future research to consider.

The PCOS survey also demonstrated **a minority of participants had used or referenced official statistics** which raises several questions; such as, is there an optimal level of statistics usage? Furthermore, if members of the public state that they 'use' official statistics, how much does an individual need to interact with the statistics in order for us to be confident that statistics is serving the public good? It is also likely that developing a better understanding of the role of intermediaries or propagators in communicating statistics would also further the understanding of what statistics serving the public good means. Additionally, exploring what the public good means to the public will be instrumental in being able to answer these important questions.

A further theme which emerged in these findings relates to **confidence working with numbers, numeracy, statistical literacy, and economic literacy**. Evidence from the Perils of Perception survey demonstrated that participants who feel less comfortable working with numbers are less accurate at estimating causes of death. This pattern was further seen in research on Coronavirus misinformation (Roozenbeek, et al., 2020) as well as evidence gathered by the CES Task Force. Together this information highlights the importance of these skills in ensuring that statistics is serving the public good. Further investigation into population levels of numeracy and statistical literacy, and how these can be improved, would be a valuable pursuit. Improving these skills would likely have far reaching positive impacts on individuals, beyond just understanding official statistics.

Similarly, evidence (Runge & Hudson, 2020) was presented showing the negative impact that 'unnecessary jargon' may have on non-expert users understanding economic concepts. Suggesting that economic concepts should be taught more widely in schools demonstrates that this is a significant task to address. In order to improve the communication of economic or statistical matters, efforts should also be made by producers and communicators to ensure that their publications can be understood and correctly interpreted by a wide range of individuals. More story-

based interpretations of statistics may be one way of communicating clearly with those who have less familiarity with statistics, as evidence from the CES Task Force suggested. However, it would also be important to consider how this can be done whilst maintaining neutrality on statistical findings and the potential need for training to do this effectively. Publishing statistics alongside commentary may respond more to the needs of users, but if the commentary appears agenda-driven this could be extremely damaging to the perceived trustworthiness of official statistics.

Economic Value

The Monetary Value of Statistics (United Nations Economic Commission for Europe, 2018)

There is clear value in statistics helping individuals to make decisions, but the value of statistics goes beyond this. The framework on the 'Recommendations for Promoting, Measuring and Communicating the Value of Official Statistics' created by the CES Task Force highlighted the importance of developing methodologies which can quantify the monetary value of statistics. This was for two reasons. Firstly, official statistics should enhance governmental policy and decision-making, therefore leading to increased prosperity. If this is the case, the prosperity generated from the use of official statistics should be quantifiable, but there is a need for methodologies to be developed which are capable of doing this. Developing these methodologies would help to demonstrate the value of statistics as serving the public good.

Secondly, with budgets decreasing and scrutiny increasing, NSOs have a pressing need to demonstrate their value. In order to show that they are using public funds in an effective way, quantifying the monetary value of statistics would be persuasive in demonstrating their importance and worth. Furthermore, it is imperative to secure adequate funding to ensure that official statistics can be of sufficient quality, or statistics producers may be forced to compromise the quality of their statistics (MacFeely, 2016). Methodologies which can quantify the monetary value of statistics are therefore valuable in supporting the argument for high quality official statistics, which can serve the public good, and economically justifying the NSOs who produce them.

Economic Statistics and Timeliness

The Bean Review (2016) highlighted the need for more timely economic statistics, stating that there is often a trade-off between timeliness and accuracy, with timeliness often being sacrificed for accuracy. Faster economic indicators which are inaccurate have the potential to be misleading but Sir Charles Bean (2007), in a speech during his time as Chief Economist of the Bank of England, made the point that it is better to have "timely, albeit inaccurate, initial estimates to factor into our decisions. But we do need to recognise that such early estimates may be revised…" (p. 2). This suggests that having indicators which are available much more quickly may be advantageous in allowing faster decisions to be made, but with the stipulation that these indicators will be subject to change as more information becomes available. In this sense, Sir Charles Bean is suggesting the agility provided by less accurate but timely data would be preferrable to the alternative but ideally methods would be developed which involved less of a compromise on the accuracy

of the data as there are inevitable risks involved with using data which may be innacurate.

Empirical research published by ESCoE investigated the costs associated with these trade-offs. Kara and Lennard (2020) estimated the cost of statistics which were not timely and accurate enough to inform fiscal and monetary policy decisions during the Global Financial Crisis. They projected that if data had been more timely and more accurate, interest rates could have been lowered more quickly. The authors suggest that this could have increased output by an estimated £12 billion, meaning 150,000 more people could have been employed at this time. This calculation allowed them to quantify the cost of inadequate data, in this case inaccurate statistics and lengthy lags in data availability, which subsequently clearly do not serve public good; these statistics have led to costs instead.

Research by Koop, McIntyre, Mitchell, and Poon (2018) similarly highlighted the potential harm caused by outdated regional economic updates. To decrease these delays, Koop and colleagues developed large econometric models to provide more timely information about regional economics. Previous regional statistics were almost two years old by the time they were applied but the econometric models Koop and colleagues developed provided much faster and more useful insights. This information is particularly important given that each region of the UK had relatively different impacts and recovery from the Global Financial Crisis therefore making more timely regional updates even more crucial and in the public's best interests.

Social Value

This section will consider how we can define and measure the social value that statistics provide. In considering how to define and measure the social value of statistics, drawing insights from other products which meet the criteria for a public good may further current understanding. One such product which has been extensively considered in terms of value is free television broadcasting provided by the British Broadcasting Corporation (BBC).

Building Public Value

Free television broadcasting meets the criteria for a public good and the BBC arguably has a responsibility to provide a national service. The BBC (2004) has written at length about the fact that it wants to go beyond the basic principles of delivering a public good to create **public value.** It defines this as providing value to individuals (to inform, educate, and entertain them), social value to citizens (as it brings social benefits to communities), and economic value (to provide insights into commercial markets), which may parallel the approaches to public good identified in this review.

The BBC maintains principles of universality, fairness, equity, and accountability, and it also aims to maximise the social value it offers by connecting people and creating a sense of belonging. It refers to this as **social capital**, which it defines as a "network of shared values, traditions and experiences that people hold in common" (p.36). By appealing to a wide range of people, the BBC has essentially created a communal factor which brings people together to inform, educate, provide insights to

commercial markets and create benefits within communities. It could be argued that this is similar to the benefits that could be achieved by statistics when it is said to serve the public good, therefore considering public value in this way may be useful.

Data Gaps

The abundance of statistics we are exposed to may make it harder to believe that there are gaps in knowledge, and that data is not completely thorough, but this still remains the case. There is very little known about how the public use data but if data or statistics do not exist, or are incomplete, this could potentially compromise statistics serving the public good. Identifying gaps in data is a concern for OSR but this is a very substantial task to undertake. Due to the potential damaging effects of data gaps, there is also a significant challenge from others; for example, the Missing Numbers blog (Powell-Smith, 2020) which contains 'The Government Data Graveyard'. This is a list of datasets which have been discontinued and "So far, this website knows of about 92 datasets".

There is also significant challenge to data gaps from Caroline Criado-Perez who works to identify data gaps which exist in regard to women. In her book, Invisible Women (2019), she states "When your big data is corrupted by big silences, the truths you get are half-truths, at best" (p. 2). Criado-Perez has identified a consistent gap in data which persists wherein data collection tends to focus more on men as they are considered 'the default'. If data is not being collected in an equitable and fair way, this represents a serious problem as statistics cannot serve the public good if they only reference the data of men.

It is important to state that there can be different explanations for these gaps in data and that discontinuing the collection of certain data may not be inherently problematic. The Code of Practice for Statistics actually suggests that statistics producers review their statistics periodically to consider whether statistics should be continued, discontinued or changed (in discussion with users) to ensure that the statistics being produced are relevant to users. However, while consistent gaps in sex data may point towards systemic oversights, other data may prove elusive to researchers due to the challenges of capturing some types of information. For example, a report into the measurement of the numbers of vulnerable children in the UK by the Children's Commissioner (2017) outlined the difficulties associated with measuring this number accurately. When it is clearly in the interests of the public good for a data gap to be filled, for instance in the case of measuring the numbers of vulnerable children, innovative approaches and methodology potentially create paths of progress for researchers to consider in the most difficult of cases.

Public Good of Statistics Research Programme

This literature review had the aim of identifying current approaches to the concept of the public good of statistics, to understand what is known and what is not known. This review of the approaches to the public good of statistics has demonstrated that there are different ways of conceptualising the public good of statistics; it can be considered in terms of how the public perceive statistics (how trustworthy they are) and how useful they are.

Firstly, the legislative approach to the public good of statistics was reviewed as this demonstrated that statistics serving the public good has a clear basis in legislation which has only been further supported by the introduction of the Digital Economy Act (2017). The definition of statistics serving the public good is very clearly described but it is important that work continues in trying to understand precisely how this can be achieved so that both the public and policy makers have a clear understanding of what the public good of statistics is and benefit as much as possible from official statistics.

Secondly, the empirical research approach demonstrated that there are still many unanswered questions and much scope for empirical research to better understand how best statistics can serve the public. The empirical research approach demonstrated some consistent themes around the trustworthiness of statistics and statistics producers and numeracy and statistical literacy and which will help to guide future research on this topic. It also uncovered the fact that there is still further research needed to understand exactly what statistics serving the public good means for the public themselves.

Thirdly, the economic value of statistics was considered. This approach made it clear that timely and accurate statistics are in the interest of the public good as ill-informed fiscal and economic policy can be extremely costly to the public and the country as a whole. Furthermore, it is important to develop methodologies that help to quantify the value of official statistics; this is needed in part to show there is enough value to justify the budgets dedicated to NSOs, who require considerable resources to produce high quality statistics. Evidence from the CES Task Force and ESCoE demonstrates that important work is underway in developing and improving methodologies in terms of the economic value of statistics, which is very encouraging in terms of achieving the vision of statistics serving the public good.

Finally, the social value of statistics serving the public good was discussed as parallels can be drawn between what the BBC is trying to achieve in terms of social value and what OSR wants to achieve for statistics. Introducing the concept of social capital to this discussion may offer an avenue for research in this area to consider. The problems resulting from data gaps in compromising the public good of statistics was also discussed, although again it is encouraging to see the considerable external challenge to this problem which is taking place.

This literature review had the objective of taking the first step towards answering the question ‘whose needs are served by statistics?’ and it seems that some needs are

being met but there is still much work to be done to be confident that statistics are serving the public good. This information is vital in helping OSR's Public Good of Statistics Research Programme to proceed and make progress in ensuring that statistics benefit the widest range of users possible.

References

BBC. (2004). *Building Public Value: Renewing the BBC for a digital world.* Retrieved from https://downloads.bbc.co.uk/aboutthebbc/policies/pdf/bpv.pdf

Bean, C. (2007). *Risk, Uncertainty and Monetary Policy (Speech given to Dow Jones).* Retrieved from https://www.bankofengland.co.uk/-/media/boe/files/speech/2007/risk-uncertainty-and-monetary-policy

Bean, C. (2016). *Independent Review of UK Economic Statistics.* Retrieved from https://www.gov.uk/government/publications/independent-review-of-uk-economic-statistics-final-report

Children's Commissioner. (2017). *On Measuring the Number of Vulnerable Children .* Retrieved from https://www.childrenscommissioner.gov.uk/wp-content/uploads/2017/07/CCO-On-vulnerability-Overveiw-2.pdf

Criado-Perez, C. (2019). *Invisible Women.* London: Chatto & Windus.

Digital Economy Act. (2017). Retrieved from http://www.legislation.gov.uk/ukpga/2017/30/contents/enacted

Duffy, B. (2018). *The Perils of Perception: Why We're Wrong About Nearly Everything.* London: Atlantic Books.

Grenander, U., & Miller, M. I. (1994). Representations of knowledge in complex systems. *Journal of the Royal Statistical Society: Series B (Methodological), 56*(4), 549-581.

Hand, D. J. (2008). *Statistics: A Very Short Introduction.* New York: Oxford University Press.

Ipsos MORI. (2020, 11 25). *Evidence use: Under the microscope of public opinion in a time of crisis.* Retrieved from Ipsos.com: https://www.ipsos.com/ipsos-mori/en-uk/more-britons-want-government-make-fast-decisions-crisis-wait-all-evidence-long-theyre-open-about-it

Ipsos MORI. (2020). *Perils of Perception.* Retrieved from https://perils.ipsos.com/slides/index.html

Ipsos MORI. (2020). *Veracity Index.* Retrieved from https://www.ipsos.com/ipsos-mori/en-uk/ipsos-mori-veracity-index-2020-trust-in-professions

Kara, A., & Lennard, J. (2020). *Valuing Economic Statistics: A Case Study.* Economic Statistics Centre of Excellence (ESCoE).

Koop, G., McIntyre, S., Mitchell, J., & Poon, A. (2018). *Regional Output Growth in the United Kingdom: More Timely and Higher Frequency Estimates, 1970-2017.* Economic Statistics Centre of Excellence (ESCoE).

MacFeely, S. (2016). The Continuing Evolution of Official Statistics: Some Challenges and Opportunities. *Journal of Official Statistics, 32*(4), 789-810.

McDonald, L. (2012). *Florence Nightingale and Hospital Reform: Volume 16 of the Collected Works of Florence Nightingale.* Waterloo: Wilfrid Laurier University Press.

Murray, S., & Gal, I. (2002). Preparing for Diversity in Statistics Literacy; Institutional and Educational Implications. *Sixth International Conference on Teaching Statistics* (pp. 1-8). Cape Toown: IASE.

NatCen. (2019). *Public Confidence in Official Statisics.* Retrieved from http://natcen.ac.uk/our-research/research/public-confidence-in-official-statistics/

Nicholson, W. (1978). *Microeconomic Theory: Basic Principles and Extensions.* Fort Worth: The Dryden Press.

Nightingale, F. (1858). *Notes on Matters Affecting the Health, Efficiency and Hospital Administration of the British Army.*

Powell-Smith, A. (2020). Missing Numbers. Retrieved from https://missingnumbers.org/

Roozenbeek, J., Schneider, C. R., Dryhurst, S., Kerr, J., Freeman, A. L., Recchia, G., . . . van der Linden, S. (2020). Susceptibility to misinformation about COVID-19 around the world. *Royal Society Open Science*, 201199.

Runge, J., & Hudson, N. (2020). *Public understanding of economics and economic statistics.* London: Economic Statistics Centre of Excellence.

Statistics and Registration Service Act. (2007). Retrieved from http://www.legislation.gov.uk/ukpga/2007/18/contents

United Nations Economic Commission for Europe. (2012). *Making Data Meaningful; A Guide to Improving Statistical Literacy.* Geneva: United Nations.

United Nations Economic Commission for Europe. (2018). *Recommendations for Promoting, Measuring and Communicating the Value of Official Statistics.* New York & Geneva: United Nations.

Waind, E. (2020). *Trust, Security and Public Interest: Striking the Balance.* ADR UK Strategic Hub.

Wallman, K. K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the Americal Statistical Association, 88*(421), 1-8.

Winton Centre. (2020, 12 01). *What has been happening to attitudes to government, science and COVID-19 in the UK during the pandemic?* Retrieved from

Winton Centre for Risk and Evidence Communication: https://wintoncentre.maths.cam.ac.uk/coronavirus/what-has-been-happening-attitudes-government-science-and-covid-19-uk-during-pandemic/