Contemporary Social Science

Themed Issue on 'Evidence Policy in a Digital Society'

Guest editors: Linda Hantrais (corresponding guest editor) and Ashley Thomas Lenihan

Opportunities and challenges for official statistics in a digital society

Paul Allin (Corresponding author), Department of Mathematics, Imperial College London, London, Room 547 Huxley Building, South Kensington Campus, London SW7 2AZ, UK, p.allin@imperial.ac.uk

Notes on contributors

Paul Allin is a Visiting Professor in Statistics in the Department of Mathematics at Imperial College London. His research interests are the measurement of national wellbeing and progress, and the use of these measures in politics, policy, business and everyday life. He lectures on official statistics and chairs the UK Statistics User Forum and the Advisory Panel of the What Works Centre for Wellbeing. He spent 40 years as a statistician, researcher and policy analyst in UK government departments and agencies, including directing the Measuring National Wellbeing programme at the UK Office for National Statistics (ONS).

Disclosure statement

No potential conflict of interest was reported by the author. The author gratefully acknowledges advice provided and extensive material published by ONS. Thanks also to David Hand, two reviewers and the guest editors for many helpful suggestions.

Views expressed in the paper are the personal views of the author and do not necessarily

represent those of ONS, the UK Statistics Authority, the Statistics User Forum or the What Works Centre for Wellbeing.

Opportunities and challenges for official statistics in a digital society

Abstract

National systems of official statistics are expected to provide governments, businesses and the public with data about the economic, demographic, social and environmental situation. Digitisation of data collection for official statistics is mooted as having a major potential impact on society. Greater use of administrative data held in government is proceeding slowly. Evidence is limited, but access to big data − from satellites, point-of-sale systems and social media − is being explored and trialled. Initiatives are driven by new requirements for official statistics as well as pressures on traditional data collection from households and businesses. Effective government use of new data sources has the possibility of creating the ultimate evidence base for policies intended to improve lives, and anonymised data are also being made available to other researchers. The article uses the UK as a case study to outline how the big data evidence-to-policy process is intended to work in that context, and to assess the challenges faced in making it work as intended. We conclude that technical developments need to be accompanied throughout with attention to the marketing of official statistics and engagement with all users and potential users. The goal is to produce trusted as well as trustworthy statistics.

Keywords: UN fundamental principles; evidence-based policy; big data; administrative data; user requirements; research ethics

Introduction

Official statistics are an established part of the evidence base for society. They are

envisioned, in the fundamental principles for official statistics promoted through the United Nations (UN), as ‘an indispensable element in the information system of a democratic society’, designed to provide governments, businesses and the public with ‘data about the economic, demographic, social and environmental situation’ (UN Statistics Division, 2014, pp. 1–2).

At first glance, official statistics are embedded in digital society: almost every country has a national statistics website listed on the UN’s webpages (at https://unstats.un.org/home/nso_sites/), with outputs increasingly digital by default and with innovative ways of publishing data and statistics. Statisticians are using social media to draw attention to newly-published summaries and to alert users to developments and formal consultations, for which responses are gathered online. However, they appear to be using new digital media mainly to replicate traditional means of consulting known stakeholders, albeit with obvious improvements in efficiency, rather than any concerted attempt to widen engagement and contact new users. Messages are often broadcast rather than seeking to engage in dialogue, apart from when using new technologies to assist in interviewing survey respondents (for example Hand & Vichi, 2018).

This article argues that significant challenges exist for official statistics in the digital society, as well as major opportunities. The main challenge is for official statistics to remain useful and relevant: according to the UN’s fundamental principles, official statistics must always have ‘practical utility’. Codes of practice seek to put the test of practical utility into effect. We propose that official statisticians should also make greater use of well-known marketing ideas as they engage with digital society to make

statistics as useful as possible in the contemporary information space, where potential users can readily resort to alternative sources of evidence. This process is about building trust in official statistics as well as producing statistics that are intended to be trustworthy through their technical qualities.

In terms of opportunities, we explore how official statisticians are beginning to make greater use of external big data sources, building on their use of administrative data held by governments, in providing relevant official statistics. The term ‘big data’ is difficult to define precisely, although a significant attribute of a set of big data is that it is treated as a complete (not a sample) set of a very large number of people, transactions or observations. While administrative datasets often act as the testing ground for processes and procedures for accessing and analysing new sources (including recognising that big data may be incomplete), official statisticians are also now looking at the potential of sources such as satellite imaging, mobile phone meta-data (data about phone usage rather than the content of conversations), transaction data (such as point-of-sale prices) and social media.

We then illustrate these opportunities and challenges with a case study examining recent UK legislation to understand the goals of new data collection processes for compiling official statistics. Specifically, we explore the Office for National Statistics’ (ONS) current plans for a census in 2021 that paves the way for administrative data censuses thereafter, and then discuss how big data sources more widely are beginning to be harvested for UK statistics.

Using official statistics as evidence

Although the fundamental principles referred to above suggest that official statistics are equally available for all in society, in reality there are practices, procedures and cultures that can result in government users taking priority. Predating consolidation of the current official statistics system, during the second half of the twentieth century, there is a history over many centuries of a need for information about the state, for use by state authorities and funded by them. Data sources for official statistics still include administrative systems of government. Codes of practice, statutory underpinnings and other ways of ensuring that the fundamental principles now apply across the board therefore need to recognise the complexities of how official statistics are used and from where the data are drawn. For example, as well as feeding directly into government policymaking, the same statistics should be available for people outside of government to assess its performance.

Moreover, while official statistics are understood to be used in many situations, including commercial decision-making, appraisal of policy options, and within educational curricula, we know little in detail about how they are used, or by whom. Descriptions of the user base are invariably a list of users known to producers, along with some indication of the intended categories of users whom producers aspire to serve. The UK Statistical Authority's (UKSA) report on its assessment of the state of UK National Accounts is typical: 'Business, research and education communities, the media and the public use the National Accounts to provide a basis for analysing the UK's economic performance … [including to investigate] causal mechanisms at work within the economy' (UKSA, 2015, §§2.12). One hint of greater specificity is that 'Large parts of the National Accounts have legal standing in the European Union'

(UKSA, 2015, §§2.14). Vichi, Valente Rosa & Ruane (2015, p. 4) construct a schematic table of type of users to distinguish between institutional and non-institutional users, as well as allowing for heavy, light or non-usage, and whether institutional users have a general, specific or research interest. However, apart again from European institutions with a statutory interest in European-wide statistics, no detail is available on other users, nor any indication of the relative size of each user category.

Users have many sources that they can, and do, call on to form their overall evidence base, for example for assessing a country's economic performance, which specific official statistics are designed to measure. Policymakers, commercial decision-makers, journalists, analysts and the public will draw on a range of facts, opinions, personal and other experiences to assess the state of the economy. The senior UK government minister reported as saying 'Never mind the figures, the economy is "healing"' (Helm & Wood, 2012) may or not have had better or more up-to-date information than from the official statistics but he was giving weight to views from the business community. Likewise, the central bank draws not only on official data but also on its own surveys and contacts in coming to a decision on monetary policy (Bank of England, 2019).

Undoubtedly, more information and more easily accessible information is available in the digital society than in the days when official statistics were published on paper and reported in a small number of media outlets. Livingstone, Blum-Ross, Pavlick & Ólafsson (2018, p. 1) argue that 'digital media sit alongside' established ways of interacting, playing and communicating within households, rather than 'displacing them'. This is also happening within public policymaking. Mayer (2013, p. 2), a member of the independent committee advising the UK Government on natural capital,

has observed that, while measurement is critical, 'it does not in itself automatically translate into policy'. Policymaking has inputs other than official statistics, with no guarantee that policymakers will access all relevant statistics, whatever the extent of digitisation. The gold standard can be understood to mean that successful policy outcomes depend on the development and use of a sound evidence base, understanding and managing the political context, and focussing on delivery from the outset. Clear opportunities can be found for using statistical evidence throughout the policy process.

The caveat, though, is that policy analysis is an imprecise science. Manski (2013, p. 3) has provided many examples of the value of moving policy analysis 'away from incredible certitude and towards honest portrayal of partial knowledge'. If policymaking is informed by evidence rather than based solely on evidence, then this places a responsibility on policymakers to get the balance right, and on politicians, auditors and the public to form a view on the extent to which relevant evidence has been used reasonably. Rutter (2012, p. 7) reports on 'positive developments' in evidence-based policymaking, while noting that 'there remains a gap between aspiration and practice' and summarises how use of evidence and evaluation might be better embedded in the policymaking system. One such development is the formation of centres for evidence-informed policy and practice in the UK, elsewhere in the EU, and in the US and Australia, which are known in the UK as 'what works' centres.

The future role of official statistics is far from given. We have not dwelt here on potential cultural and other barriers to greater adoption of policy processes and public discourse based on, or informed by, systematic evidence. For example, a case can be

made for improving the skills of politicians, journalists, decision-makers and the public in finding, understanding and using official statistics.

Taking a marketing approach in official statistics

The production of official statistics might reasonably start with identifying users and their requirements if the goal of practical utility is to be met. In practice, determining user requirements is seldom done with any precision. Rather, the ethos of official statistics appears to be to provide a body of data to match broadly the scope of the fundamental principles for official statistics, and then to work with users and user groups to meet needs using available statistics.

To meet the full vision for official statistics, all official statistics could be considered as products to be subject to a thorough marketing regime. This would require not just promotion and making improvements in availability, but also understanding the need for data and how they may fit in, or even shape, a market. Then the issue would be how to provide better answers to specific questions, by finding out what the questions are beforehand, and collecting pertinent rather than general data.

One obvious objection to marketing is that official statistics are not market products but public goods, invariably published free of charge. Official statistics offices may respond to specific data requests, for free if they involve marginal effort, otherwise in exchange for a fee. But these requests often depend on users knowing how to invoke such a service.

To move to a more open approach, it is helpful to consider what Collins (2010) originally posed as four questions to developers of any product or service. The first question, echoing the need for utility in official statistics, asks simply: What are customers' needs? To answer this question calls for active and on-going user engagement and consultation, with iteration between producers and users. This process does not generally seem to be undertaken in the production of official statistics, at least beyond interacting with key users. A benefit of engaging with users is that the inevitable trade-offs that have to be faced between technical aspects of quality – such as accuracy, timeliness, and level of detail – can then be made with at least some understanding of how they will affect the overall utility of the statistics. The case for a fresh focus on the full use of official statistics, along with examples of positive developments and topics still needing more work, has been made by the Statistics User Forum (2019, GOS0024) in evidence to a Parliamentary inquiry into UK official statistics.

The second two of Collins' marketing questions are about competition and understanding the market. National statistics offices do not have a monopoly on publication of data and statistics. The need to understand the competition and to brand official statistics clearly and distinctively has become more important. Market research should help not only with understanding requirements but also how best to reach potential users.

The final marketing question is about the value proposition of the product. Official statistics have value to society that should be recognised, alongside consideration of costs involved in collecting and processing data. The United Nations Economic Commission for Europe (UNECE, 2018) has highlighted the value of having official

statistics. It recommends ways for statistical offices to promote, measure and communicate this value. UNECE (2019) reports that seven countries are pilot-testing the recommendations, to provide further methodological guidance, while noting that the 'few' attempts at calculating the monetary value of official statistics so far 'have demonstrated that Official Statistics bring net benefits'.

We are not, of course, asserting that user engagement is missing from official statistics, or that adopting more of a marketing approach will be the panacea. There are many cases of good practice on which to build. To take as an example, an assessment of the UK National Accounts statistics (UKSA, 2015, §§1.3.1) concluded that ONS has improved its engagement practices, but qualified this by adding 'at least with key stakeholders'. UKSA (2015, §§1.3.3) also notes that users 'generally accept ONS's priorities of developing the National Accounts against international standards over other potential enhancements [but] that ONS seems slow to modify methods used to compile the National Accounts in response to changes in the structure of the UK economy'. The case for better engagement is clear from the 340 assessments undertaken by the UK regulator of official statistics, where 250 assessments (i.e. almost 3 in 4) included a requirement to improve engagement (Statistics User Forum, 2019, GOS0030).

The need for big data in official statistics

We identify three broad pressures on existing sources of data for official statistics that are giving rise to the search for new, big data sources. First, increasing requirements from users. The task is to build an evidence base for policy and for wider use that increases significantly the availability of 'quality, accessible, timely and reliable disaggregated data' (for example UN, 2015, §48). The defining feature of many

emerging requirements is that statistics should be able to be highly disaggregated, to allow sub-sectors of the population of people or resources at risk to be monitored. Second, the willingness of businesses and households to respond to official surveys is diminishing. Third, official statistical organisations are invariably funded entirely, or very largely, out of governmental budgets, including in some cases as part of international development aid, and these budgets are under pressure.

These pressures can be tackled in many ways, not least by making the case for the value of official statistics, as touched on above. We concentrate in this article on the opportunity to cut costs and improve the timeliness of official statistics through data collection that embraces the digital society and makes much greater use of existing big data sources, including official administrative data. At the global level, the UN (2015, Goal 17.18 and §76) has called for new data sources for official statistics to be taken forward especially through 'appropriate public–private cooperation to exploit the contribution to be made by a wide range of data, including earth observation and geo-spatial information'. This does not so much constitute a data revolution as a change in the mix of types of data − censuses and surveys, data from administrative systems, big data from non-official, transactional or observational systems – with data joined-up in ways that preserve the confidentiality of individual people and businesses.

Developing the use of digital data in UK official statistics

While the UK has a relatively sophisticated system of official statistics, it is far from immune from the need to find new data sources. Documents released by ONS (2015), reveal how UK official survey response rates fell markedly during previous decades (and this is a global phenomenon). For example, in one of ONS's major household

surveys the response rate fell year-on-year from 79% in 1993 to under 48% in 2014, the latest figure then available. Gaining a response of under a half of all respondents selected is bound to shake the confidence of the survey team and their users, despite all the rigour in constructing representative samples and identifying the characteristics of non-respondents, where possible. Not surprisingly, when introducing a strategy for UK official statistics shortly ahead of the detailed evidence on response rates being published, the National Statistician declared 'I fully expect that, in five years' time, what we will be doing will be radically different. More data will be real-time; our services will be digital by default' (UKSA, 2014).

In the UK, the trajectory towards new data sources is building on use of administrative data. Historically, UK statistics derived from administrative data have been designated as management information, for the purposes of supporting delivery of that particular public service. However, in some cases, administrative data have also been mined through secondary analysis, as part of the evidence base provided by official statistics, with unemployment and crime as two of the major examples. A report by the Royal Statistical Society recognised that administrative data had a role in understanding unemployment and labour market trends, but that 'the monthly headline count should be based on the Labour Force Survey', using an internationally agreed definition of the unemployed (Bartholomew, Moore, Smith & Allin, 1995). This recommendation appears not to have been revisited even though the LFS is the household survey that was given as the example of declining response rates above. Each month, over the last 40 years, an administrative system has been mined to produce 'claimant count' statistics of characteristics of up to 3 million people. The position on crime statistics is more nuanced, with at least some administrative data (for example, on infrequent but violent

crime and on fraud) needed to supplement survey data on households' experience of crime.

The 2001 census in England and Wales was the first in which a strategy for the development and release of results was included in the census preparatory White Paper as government policy. At its heart was the balancing of cost and value, that 'the investment of time and resources in a national census is only justified if the results are made accessible to users speedily and in a clear and usable form'. To this end, 'technological developments should be harnessed in order to improve the accuracy, timeliness, accessibility and user-friendliness of published output' (ONS, n.d.). An external service provider processed the census (ONS, 2013). Digitisation was employed and over 24 million census forms were scanned, with responses coded electronically using automated and computer-assisted methods. In the previous census, manual coding of occupation and industry was undertaken and restricted to 10% sample of forms.

The greater use of administrative data in building an evidence base is not just an issue for the compilation of government statistics and research. The academic community, long used to accessing de-identified survey data for secondary analysis, has also recognised the value of secondary analysis of administrative data. An Administrative Data Research Network (ADRN) was formed in 2013, following a report arguing that administrative data in research need to be used more effectively across the UK 'for public and policy benefit' (Administrative Data Taskforce, 2012, p. iv). The network noted that 'the UK is lagging behind fellow European countries and other parts of the developed world' (ADRN, n.d.).

Differences in approach and culture can be found between the UK government and the devolved administrations. However, use of administrative data for statistics and research has generally proceeded only slowly across the UK. It was apparent by 2017 that better legislation was necessary, with an opportunity for this issue to be included in what became the Digital Economy Act (UK Parliament, 2017).

Writing to the Minister responsible, the head of the UKSA noted that:

> The existing legal framework governing access to data for official statistics is complex and time consuming. The proposals in the Bill, by making use of data already held across government and beyond, would deliver better access to administrative data for the purposes of statistics and research, delivering significant efficiencies and savings for individuals, households and businesses. (Dilnot, 2016, p. 1)

Chapter 7 of the Digital Economy Act 2017 amends legislation from 2007 to allow for the disclosure of information by UK public authorities, including HM Revenue and Customs, to UKSA to compile statistics. The Act thus gives ONS a right of access to information held by government departments, other public bodies, charities and large and medium-sized businesses for statistics and research purposes. The legislation also provides, though chapter 5 of the Act, for the sharing of de-identified administrative information by public authorities with other persons for the purpose of research. This arrangement is overseen by UKSA, with a code of practice and processes to maintain confidentiality of data subjects. These two data-sharing provisions are included with many other aspects of the digital economy in the legislation.

Although progress with data-sharing provisions was keenly followed by the statistics and research community as the draft legislation made its way through Parliament, these provisions did not attract attention from the media and the public. The lack of public interest may suggest that other aspects of the legislation, such as broadband speeds, regulation of social media platforms and funding of the BBC, were more controversial. The Act encouraged ONS to do all it could to reassure businesses and households that their personal data are secure and are not disclosive when the statistics derived from the data are published. The continued privacy of individuals is 'absolutely critical to maintaining public trust', in the words of one witness examined by a Parliamentary committee. Another witness stressed the value of administrative data for statistical purposes and the need for businesses and households to be spared the time and effort of filling in survey forms (UK Parliament, 2017).

Ahead of the revision to the legislation, ONS had been planning for the next decennial census in England and Wales, in 2021. This will predominantly use online data collection from households and will be a stepping stone for 'administrative data' censuses thereafter. This means that censuses from 2021 onwards will require, according to ONS planning documents, a combination of record-level administrative data held by government, a population coverage survey, a population characteristics survey and commercial and other non-survey data sources. Evidence of feasibility is emerging from ONS's annual assessments of its progress towards administrative data censuses. The 2017 assessment (ONS, 2017a) was mainly buoyed up by the passing of the legislation, along with technical work on the ability to link data efficiently and accurately. Other criteria, repeated from the 2016 assessment, include the extent to which user needs for statistical information, say on the distribution of income as well as

the demographic features of the population, are likely to be met, value for money, and acceptability to stakeholders, including the public.

ONS's 2018 assessment, the latest available at the time of writing, re-iterated the broad aims of this work, including that decision-makers need census-type information much more frequently than every 10 years. In particular, they 'need to understand the changing nature of our population and migration to and from the UK more regularly and locally' (ONS, 2018, Section 1). The assessment records progress in developing methodology, working with data owners and publishing results of research into the feasibility of using new sources. However, the formal assessment has hardly changed from that of the previous two years, leaving much to be achieved between 2018 and 2023, when ONS will 'make a recommendation on the future of the census'. ONS's assessment of where they will be in 2023 is looking hopeful, with greatest uncertainty over their ability to meet the needs of users for information on population and household characteristics, rather than just providing estimates of population and household numbers (ONS, 2018, Section 2).

One specific development is that data on mobile phone usage appear promising as a source of information on commuting patterns (ONS, 2017b). ONS used a data analytics company to model the number of flows between local authority districts, derived from data from one mobile phone network operator. This illustrates some general aspects regarding the use of other kinds of data, including around access and data protection, that need to be negotiated. It also raises questions about the transparency of methodology. In its experimental work, ONS (2017b, Annex A) reported that the methodology 'is owned by the mobile network operator and the data analytics

company', indicating a possible trade-off between aspects of the UN fundamental principles, in which transparency is sought.

The transformation of the census is the most high-profile UK official statistical work to make greater use of big data and is acting as the focal point for other developments. In a progress report on the strategic development of UK official statistics (UKSA, 2017), 'unlocking data' through the legislative change was noted as an achievement, though 'much better use of data and clearer assurance for the public on how it is being used' was still waiting for the Digital Economy Act to come into force the following year. Moving several business surveys to online data collection was another major achievement. ONS opened a data science campus in 2017 to explore new data sources and techniques (https://datasciencecampus.ons.gov.uk/). More is promised, including monitoring UK progress towards the UN sustainable development goals. While what is sometimes called a data revolution is clearly underway in the UK, development plans stretch forward into the next decade.

Challenges and implications for official statistics

This article has focussed on the UK as a case study because the UK, like many other countries with official statistics based mainly on censuses and surveys, is entering the digital society from a different position to countries with register-based systems. The fundamental principles for official statistics encourage international cooperation and exchange of good practice. Statistics Netherlands (2014, pp. 153–160) has published guidelines on quality assurance for administrative data and registers and also has a centre for big data statistics. Young, Hyman & Rater (2018, p. 337) reported how big data sources can be used to compile a register of agricultural land holdings within a city,

concluding that ‘Web scraping, technology, and secondary data sources may be tools that are used increasingly’. In this case some of the activities of interest had no direct online presence, confirming the need to combine data from a variety of sources, including administrative data and social media.

Faced with a potential source of administrative data, crucial questions for many official statisticians are around technical quality, timeliness and reliability of the data. Hand (2018, p. 555) assessed the challenges of administrative and transaction data, urging statisticians to ‘approach the analysis of such data with the same cautious and critical eye as they approach the analysis of data from any other source’. The set of records in a database ‘may not be representative of the population to which one wishes to make an inference’, for example, or the level of detail may turn out to be inadequate for all potential analyses (Hand, 2018, pp. 557−558).

However, looking for big data sources that are of high quality, timely and reliable, falls into the trap of looking at quality in a limited way, focussing just on sources rather than also on uses. Buried deep within Hand’s 15 challenges concerning big data (number 9) is to ‘Explore how suitable the administrative data are for answering the questions’ (Hand, 2018, p. 569). Bean (2016) recommends, among other ways of improving UK economic statistics, that ONS should refocus its culture ‘towards better meeting user needs. ... Staff should be proactive, rather than reactive, in engaging with users and responsive to those users’ needs’. In this respect, the culture in Statistics Netherlands offers some lessons, since relationships, communication and agreements with users form the opening part of its quality assurance framework (Statistics Netherlands, 2014, pp. 31–40).

Arguably, the identification of users' needs should not be limited to the needs of policymakers, who are expected to draw on evidence before deciding on a suitable set of policy outputs, ranging from regulation to softer ways of influencing behaviour. Policy outputs are all essentially top-down. To make real change in society, top-down policies need to meet with bottom-up willingness by companies, civil society and local government to address the same issues. One goal of the official statistics system should be that all parties are informed by the same evidence base, hopefully leading to a common understanding and description of the issues. This approach does, however, give rise to an extensive web of inter-connected requirements for official statistics.

Well-grounded models are lacking for how official statistics, and evidence more generally, feature in policymaking and other societal decision-making. Bowen & Zwi's (2005) proposal for how evidence should be used sets out an '"evidence-informed policy and practice pathway" [that] can help both researchers and policy actors navigate the use of evidence'. The pathway envisages three stages: sourcing the evidence; using evidence; and implementing the evidence. The pathway also 'involves decision-making factors and a process which we have termed "adopt, adapt, and act"' (Bowen & Zwi, 2005, p. 0600). This kind of approach has been adopted by various 'what works' centres, which operate alongside policymakers and encourage the use of relevant evidence, including by acting as conduits for evidence providers to connect with policy users. What is less clear, however, is whether or not this has led to a significant take-up in the use of evidence. In the UK case study, official statisticians are engaging with users, as is clear from the annual assessment reports. ONS assesses its ability to move to

an ‘administrative data’ census using eight criteria, four of which are concerned with its ability to meet broad categories of information needs of users.

A marketing approach is also about understanding the competition, including through market research. Competition for the provision of statistics has been widened as users of every kind can, and often do, draw on other sources of information to substitute as evidence. The suppliers of these other sources may well be commercial and pay greater attention to marketing their material.

Digital technologies are often described as disruptive to existing business models and ways of working, so a monopoly source of data may be forced to open up by new channels, or various value-added services may turn the data into new forms of evidence. The digital society increasingly means ‘self-service’ when people want information. ONS recognises ‘inquiring citizens’ as one category of visitors to its website, and is aiming to make its website accessible to this segment of users, as well as to standard data users and analysts, who, along with ‘policy influencers’, often re-use ONS statistics in their own, online reports, which creates another form of competition. The recent UK legislation facilitating ONS access to new data sources also provides competition, in the sense that it provides for secure secondary analysis of de-identified administrative data by other authorised researchers. Both of these routes to data are ultimately under the wing of the UKSA, which oversees ONS and, in addition, has statutory responsibility for external access to de-identified administrative data. UKSA is expected to exercise these powers in ways that expand, rather than limit, the market in evidence.

Marketing calls for understanding and setting the value proposition of a product, which appears trivial, if official statistics continue to be made available free at the point of delivery. However, a case still needs to be made for recognising the value of statistics 'that support society's needs for information' (Office for Statistics Regulation, 2018, p. 15).

The marketing of official statistics should generate greater understanding of how they feature, and stand out, in the market place for evidence. Moving forward to make greater use of new data sources is not simply about finding the right key to unlock the sources. Even with an enhanced legislative base, it takes ONS months if not years to negotiate access and to turn new data into published statistics. The real-time statistics envisaged by the UK National Statistician may still be published with a lag from their reference period, while other providers might possibly enter the market sooner, with real-time statistics not assured to the quality standards of official statistics. Information may be the life-blood of a market economy, but the market in information itself raises questions about the skills and competencies in handling information, and how these skills can be honed.

One further issue to be considered in understanding the market for evidence is trust: 'confidence in the people and organisations that produce statistics and data' (Office for Statistics Regulation, 2018, p. 14). Trust is sometimes seen as the 'unique selling point' for official statistics. Recent evidence suggests that UK official statistics are generally trusted and that public confidence in the statistics is being maintained (Morgan & Cant, 2019). However, this evidence was compiled ahead of any substantive use of personal data from commercial organisations in the production of official statistics. To date,

ONS's only foray into big data to receive media attention was its exploration of the potential use of web-scraped prices and point-of-sale-scanner-price datasets. The application here is for experimental measures of consumer inflation, rather than to access sources containing personal data. Only recently has ONS started to publish experimental statistics drawn from data about where mobile phones are used, and we are some way off regular statistics from commercially-held sets of personal data. How this plays out in terms of trust could well be influenced by developments beyond the field of official statistics, for example increasing use of information from smartphones to fuel new applications, in which privacy might be eroded.

One way of building trust is for researchers to operate within a set of ethics that are designed to demonstrate trustworthiness. A version of research ethics suitable for the digital society has been published by the intergovernmental Organisation for Economic Co-operation and Development (OECD) and proposes that 'Data should be shared as openly as is feasible within the relevant legal and ethical constraints' (OECD, 2016, p. 6). But what are these constraints, who decides them, and do the people whose data are being used know and accept them? The OECD recommendations cover the ethical collection, sharing, and research use of new forms of data. They embrace mechanisms for the safe and responsible sharing of personal data, including mechanisms for the protection of privacy of data subjects, as well as for public input and accountability, all of which should be established and made public by data owners/controllers. Whether this will be done remains an open question, even though to do so would seem to be important in gaining public support for research and statistics in the digital age.

Conclusion

This article has drawn attention to the role of official statistics in a digital society at a time when developments, at least in the UK, are at an early stage. Official statistics are intended to be an important part of digital society, and their compilation and dissemination are increasingly undertaken using advanced technologies. However, it is not yet clear that digital technology is helping to make a better evidence base or better use of the evidence base. There are challenges as well as opportunities. The challenges were already present but may well be accentuated by digital technology applied to new sources of personal and commercial data. National statistics offices are not just suppliers of statistics; they should also be providing answers to questions raised in society, recognising that they have to compete for users' attention. By succeeding in doing so, they can confirm the value of official statistics to society.

Big data are expected to provide imaginative solutions, with potential data sources and technologies way beyond most of those currently within the scope of national statistical offices. For example, in July 2018, Maynooth University, Ireland, posted a vacancy for a doctoral studentship in the area of 'drone analytics for environmental monitoring'. The research programme involved 'analysing ground truth, drone and satellite data for monitoring various environmental variables in Dublin bay'. This is clearly the kind of new data source that is needed, but it is dealing with data that are far from ready for routine use in the compilation of official statistics. Evidence from research into the viability and utility of new sources of data will need transmitting along the kind of pathway envisaged by Bowen & Zwi (2005) before being implemented through an evidence-gathering environmental policy.

Given the appetite for evidence and new demands, official statistics should strive to remain driven by their fundamental principles as an indispensable element in the information system of a democratic society. Developing digital technologies and delivering new products is only part of what will be needed. The success of these developments, we suggest, will ultimately depend on the extent to which the producers of official statistics engage with users, including those in the media, policy and politics, as well as the public. In doing so, three broad areas must be addressed if official statistics are to play their full role: producing official statistics of value to society that are trustworthy and have assured quality.

Taking more of a marketing approach, as well as the more familiar focus on data quality and methodology, might facilitate all three areas. Demands for data and statistics are ever growing. Effective use of new data sources has the potential to create the ultimate 'evidence base' for policies intended to improve lives. New data sources and statistics may be necessary, but they are not sufficient to transform how we live our lives, eliminate poverty, preserve the planet and leave no-one behind in our economic and social development. To do all of this calls for, among other things, much greater integration of official statistics in decision-making, not just in public policy but also in business and in society overall. Official statistics are not just for policy, with its top-down approach to behaviour change, they are also there to help citizens and businesses take action while holding governments to account.

References

Administrative Data Taskforce (2012). *Improving access for research and policy*.

Retrieved from https://esrc.ukri.org/files/research/administrative-data-taskforce-adt/improving-access-for-research-and-policy/

ADRN (n.d.). *History*. Retrieved from https://adrn.ac.uk/about/background/history/

Bank of England (2019). *Monetary Policy Summary and minutes of the Monetary Policy Committee meeting, 1st May 2019.* Retrieved from https://www.bankofengland.co.uk/monetary-policy-summary-and-minutes/2019/may-2019

Bartholomew, D., Moore, P., Smith, F., & Allin, P. (1995). The measurement of unemployment in the UK. *Journal of the Royal Statistical Society, Series A, 158*(3), 363–417. Retrieved from https://doi.org/10.2307/2983439

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

Bowen, S., & Zwi, A.B. (2005.) Pathways to 'evidence-informed' policy and practice: A framework for action. *PLoS Med 2*(7): e166. Retrieved from https://doi.org/10.1371/journal.pmed.0020166

Collins, M. (2010). *Building a better mousetrap isn't enough*. Retrieved from https://salesandmarketing.com/article/building-better-mousetrap-isnt-enough

Dilnot, A. (2016). Letter from Sir Andrew Dilnot to Chris Skidmore MP regarding the Digital Economy Bill. Retrieved from https://www.statisticsauthority.gov.uk/correspondence/digital-economy-bill/

Hand, D.J. (2018). Statistical challenges of administrative and transaction data. *Journal of the Royal Statistical Society, Series A, 181*(3), 555–605. Retrieved from https://doi.org/10.1111/rssa.12315

Hand, D.J., & Vichi, M. (2018). The role of communication in Statistical science and

the strategies of communication for statistics users. European Statistical Advisory Committee document 2018/28. Retrieved from https://ec.europa.eu/eurostat/web/european-statistical-advisory-committee-esac

Helm, T. & Wood, Z. (2012). Never mind the figures, the economy is 'healing' says minister. *The Observer*,7 October. Retrieved from https://www.theguardian.com/politics/2012/oct/07/tories-green-shoots-speech

Livingstone, S., Blum-Ross, A., Pavlick, J., & Ólafsson, K. (2018). *In the digital home, how do parents support their children and who supports them? Parenting for a Digital Future: Survey Report 1*. Retrieved from http://www.lse.ac.uk/media-and-communications/assets/documents/research/preparing-for-a-digital-future/P4DF-Survey-Report-1-In-the-digital-home.pdf

Manski, C.F. (2013), *Public policy in an uncertain world: analysis and decisions.* Cambridge: Harvard University Press

Mayer, C. (2013). *Unnatural capital accounting*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/516947/ncc-discussion-paper-unnatural-capital-accounting.pdf

Morgan, H., & Cant, J. (2019). *Public confidence in official statistics – 2018*. London: NatCen Social Research. Retrieved from http://www.natcen.ac.uk/our-research/research/public-confidence-in-official-statistics

OECD. (2016). *Research ethics and new forms of data for social and economic research*. OECD Science, Technology and Industry Policy Papers, No. 34). Paris: OECD. Retrieved from http://dx.doi.org/10.1787/5jln7vnpxs32-en

Office for Statistics Regulation (2018). *Code of practice for statistics: Ensuring official statistics serve the public*. Edition 2.0. Retrieved from https://www.statisticsauthority.gov.uk/code-of-practice/

ONS. (2013). *Census 2001 review and evaluation. processing: Executive summary*. Retrieved from https://webarchive.nationalarchives.gov.uk/20160115133239/http://www.ons.gov.uk/ons/guide-method/census/census-2001/design-and-conduct/review-and-evaluation/evaluation-reports/processing/index.html

ONS. (2015). *Survey response rates (freedom of information documents, see document 1.1).* Retrieved from https://www.ons.gov.uk/aboutus/transparencyandgovernance/freedomofinformationfoi/surveyresponserates

ONS. (2017a). *Annual assessment of ONS's progress towards an administrative data census post-2021*. 20 June. Retrieved from https://www.ons.gov.uk/census/censustransformationprogramme/administrativedatacensusproject/administrativedatacensusannualassessments/annualassessmentofonssprogresstowardsanadministrativedatacensuspost2021

ONS. (2017b). *Using mobile phone data to estimate commuting flows*. Retrieved from https://www.ons.gov.uk/census/censustransformationprogramme/administrativedatacensusproject/administrativedatacensusresearchoutputs/populationcharacteristics/researchoutputsusingmobilephonedatatoestimatecommutingflows

ONS. (2018). *Annual assessment of ONS's progress on the Administrative Data Census: July*, Retrieved from: https://www.ons.gov.uk/census/censustransformationprogramme/administrativedatacensusproject/administrativedatacensusannualassessments/annualassessmentofonssprogressontheadministrativedatacensusjuly2018

ONS. (n.d.). *Summary of the 2001 Census process.* Retrieved from https://www.ons.gov.uk/census/2001censusandearlier/aboutcensus2001/summar

yofthe2001censusprocess

Rutter, J. (2012). *Evidence and evaluation in policy making: A problem of supply or demand?* London: Institute for Government. Retrieved from https://www.instituteforgovernment.org.uk/sites/default/files/publications/evidence%20and%20evaluation%20in%20template_final_0.pdf

Statistics Netherlands (2014). *Quality guidelines 2014: Statistics Netherlands' quality assurance framework at process level*. Retrieved from http://ec.europa.eu/eurostat/documents/64157/4374310/01-Quality-Guidelines-2014-Statistics-Netherlands-Quality.pdf/292b18bc-9bfd-426d-9282-785aabc43126

Statistics User Forum (2019). Written evidence from the Statistics User Forum (GOS 24), Public Administration and Constitutional Affairs Committee, Governance of Statistics inquiry. Retrieved from https://www.parliament.uk/business/committees/committees-a-z/commons-select/public-administration-and-constitutional-affairs-committee/inquiries/parliament-2017/governance-of-statistics-inquiry-17-19/publications/

UK Parliament (2017). Digital Economy Act, 2017. Retrieved from https://services.parliament.uk/bills/2016-17/digitaleconomy.html

UKSA. (2014). *Better statistics, better decisions: Strategy for UK statistics, 2015 to 2020*. Retrieved from https://www.statisticsauthority.gov.uk/wp-content/uploads/2015/12/images-betterstatisticsbetterdecisionsstrategyfor2015to202_tcm97-44175-5.pdf

UKSA. (2015). *Assessment report 299: The UK annual and quarterly National Accounts (produced by the Office for National Statistics).* Retrieved from

https://www.statisticsauthority.gov.uk/wp-content/uploads/2015/11/images-assessmentreport299annualandquarterlynationalaccount_tcm97-44450.pdf

UKSA. (2017). *Better statistics, better decisions: Three years on.* Retrieved from https://www.statisticsauthority.gov.uk/about-the-authority/strategy-and-business-plan/

UN. (2015). *Transforming our world: The 2030 agenda for sustainable development.* Retrieved from https://sustainabledevelopment.un.org/post2015/transformingourworld

UNECE. (2018). *Recommendations for promoting, measuring and communicating the value of official statistics.* Retrieved from http://www.unece.org/index.php?id=51139

UNECE. (2019). *What are statistics worth? Determining the value of official statistics.* Retrieved from https://www.unece.org/info/media/presscurrent-press-h/statistics/2019/what-are-statistics-worth-determining-the-value-of-official-statistics/doc.html

UN Statistics Division (2014). *Fundamental principles of national official statistics.* A/RES/68/261 from 29 January 2014. Retrieved from https://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx

Vichi, M., Valente Rosa, M.J., & Ruane, F. (2015). *The users of statistics and their role in the European society*. European Statistical Advisory Committee document 2015/1175. Retrieved from https://ec.europa.eu/eurostat/web/european-statistical-advisory-committee-esac/other-documents

Young, L.J., Hyman, M., & Rater, B.R. (2018). Exploring a big data approach to building a list frame for urban agriculture: A pilot study in the city of Baltimore. *Journal of Official Statistics*, *34*(2), 323–340. Retrieved from

http://dx.doi.org/10.2478/JOS-2018-0015