

Written evidence submitted by the Royal Statistical Society (DEB 02)

Briefing on the Digital Economy Bill 2016

1. The [Digital Economy Bill](#) ^[1] sets out legislation across a wide range of areas. This note sets out the Royal Statistical Society (RSS)'s views on Part 5 of the Bill on digital government and its **proposals for data access for statistics and research** ^[2].
2. **We welcome the provisions of the Bill** as they relate to data sharing for statistics and research. For statistics, the proposals seek to streamline existing legislative arrangements in the Statistics and Registration Service Act (2007), which already provide for data sharing for statistical purposes whilst making sure privacy safeguards are in place. For research, the proposals again build on existing practice where accredited researchers have access to de-identified data under strict conditions, so their research can contribute to the evidence base for the public good. Data sharing for statistics and research differs in nature from other data sharing proposed in the Bill in that they are intended to produce aggregate data whereas the other data sharing proposals focus on individual citizens' data.
3. We believe that **the existing provisions in the Bill do not provide the public and parliament with the assurances they need on the data sharing process**. Engagement with members of the public through focus groups and deliberative research has shown qualified support for data to be used for scientific, social or statistical research, where this has an intended public benefit, and where there are sufficient privacy safeguards in place ^{[3][4][5]}. Furthermore, many respondents to the Cabinet Office's consultation indicated they would be happy with data sharing for statistics and research provided the due process is followed. The Bill proposes that the UK Statistics Authority Board should prepare and publish a statement of the principles that they will exercise on data sharing for statistical purposes. We believe those principles - including confidentiality, transparency and integrity - should be outlined on the face of the Bill to assure the public and parliamentarians of the process.
4. We strongly advocate an **amendment to the Statistics chapter of the Bill**, to ensure that administrative data is made available to ONS on a compulsory basis **by all public authorities** after it is requested for statistics purposes. Government departments (Crown Bodies) must be **required** to provide data for this purpose, as is the case in many other developed countries. If this issue is not addressed, individual departments' power to object could continue to hinder the future production of a high quality evidence base for a wide range of topics, including Census-type information on the population and critical data on the economy.
5. **We support the proposals to streamline access to de-identified data for approved research**. They enable access to de-identified data by accredited researchers. However, we are

concerned health data is presently excluded, because this forms an important basis for research in the public interest (see Appendix 2 for examples). We agree that delayed consideration of this is necessary as public consultations are yet to conclude on the National Data Guardian's recommendations for NHS data security, consent and opt outs ^[6]. However, in due course we believe there should be mechanisms in place to include health data within the scope of the arrangements, subject to the outcomes of public debate and the safeguards that are required.

Further background and information can be found in the following appendices:

Appendix 1 - Background to proposed changes to the Bill

Appendix 2 - Case studies: the use of administrative records for research

Appendix 1 – Background to our proposed changes to the Bill

Chapter 7 – Amend legislation to improve data access for statistics

1. Strengthening ONS's access to administrative data forms a central plank of improving official statistics in this country. In comparison with statistical offices in other countries such as Canada, Finland and Ireland, ONS is lagging behind best practice and needs better access to administrative data to provide more accurate, frequent and timely statistics. The existing arrangements in the the Statistics and Registration Service Act 2007 (SRSA) require secondary legislation, and have proved cumbersome and time consuming to operate. It has not always been possible to access data from government departments, which are often averse to or uncertain about sharing data. Expert reviews, including [the Independent Review of Economic Statistics](#) carried out by Professor Sir Charles Bean ^[7], and the [Eurostat Peer Review Report](#) ^[8] which assessed provisions for UK Statistics against those of other European countries, have recommended that ONS access to administrative data from government departments needs strengthening and that lack of access is a key weakness of the present UK statistical system. One example where this will yield significant cost and quality benefits is the Census, which could, in future, potentially be replaced by data already held by government.
2. Chapter 7 of the Bill proposes amendment of the SRSA to improve the existing arrangements whereby the Office for National Statistics (ONS) can gain access to administrative data from other government departments for statistical purposes whilst continuing to safeguard personal information. The Chapter does not introduce any new purposes for data sharing for

statistics, and aims to streamline data access. The Bill therefore makes provision for the disclosure of information held by public authorities to the Statistics Board where the public authority is satisfied that the Board requires that information for any one or more of the Board's functions. Furthermore, it includes the power for the Statistics Board to require the disclosure of information held by some private undertakings.

3. We believe the legislation is an important step forward. However, we believe the Bill should be strengthened by an amendment to ensure that ONS are able to direct Crown Bodies to release their data for statistical and research purposes, in the same way as other public authorities and a characteristic of the legislation in other countries. The wider goal must be to ensure that data for the production of statistics is as accessible as possible to safe processing and analysis, with privacy safeguards, for the public good.

Chapters 5-6 – Enable access for safe researchers to work with de-identified data

4. Chapters 5 and 6 of the Bill – which cover Research, and Disclosure of non-identifying information by HMRC – should help to enable a wide range of beneficial studies. Engagement with members of the public through focus groups and deliberative research have shown qualified support for data to be used for scientific, social or statistical research, where this has an intended public benefit, and where there are sufficient privacy safeguards in place ^{[3][4][5]}. For example, research using the data held by government has led to new research and statistics on graduates' employment and earnings after higher education, allowing us to understand more about how university affects young people's life chances. This was dependent on the linkage of de-identified tax data and student loan records for 260,000 students up to 10 years after graduation ^[9]. In Appendix 2 below, we provide further examples of studies that have been reliant on data linkage, and some of those that are not possible or are substantially constrained by present caution about sharing data for de-identified research.
5. We believe, as do others, that the research dimensions of the legislation will not materially raise privacy risks provided that the safeguards proposed are applied properly ^[10]. We can point to the excellent work that has taken place in Scotland to develop their Guiding Principles for Data Linkage ^[11], which support the legal, ethical and efficient use of data for studies that link data within a controlled and secure environment – this example can inform the principles for sharing for research purposes proposed in this Bill. The Bill does not propose sharing citizen identities for research purposes, and specifies that data must be de-identified prior to research use. Adherence to this chapter of the Bill should help to ensure that this type of research no longer needs separate approval in legislation, and should therefore increase the potential for social science to produce a strong public evidence base.

6. Clause 63 (2) of the Bill states that health and social care authorities are out of scope. However, Appendix 2 shows there are many examples of studies that depend on integrating health data as part of a broader study. We understand that next steps for research using NHS patient data are being considered after the National Data Guardian's review, and public consultation on data security, consent and opt-outs^[6]. This is a highly important process to address care.data's failings. In our view the Bill should make provisions to ensure health data can be included, taking into account the outcomes from the National Guardian's Review.

Additional calls for improved data access

7. The Digital Economy Bill follows a number of high profile calls for improved data sharing. For research and statistics in particular, there is the Independent Review of Economic Statistics^[7], the Science and Technology Select Committee's Big Data dilemma inquiry report^[12], the Law Commission's scoping of data sharing between public bodies^[13], the findings of the Administrative Data Taskforce^[14, 15], as well as the UK Statistics Authority recent outline of the case for change in their policy paper^[16]. Our consultation responses have previously commented on the legislative proposals, and can be read [here](#)^[17].

Appendix 2: Case studies: the use of administrative records for research

This appendix provides examples of how administrative records have been, or could be, used to address key areas of public significance.

These first three examples describe **potential research based on administrative records which has not been possible as a result of the existing arrangements in legislation.**

- **Investigating what affects winter mortality**

Research requested by the Chief Medical Officer for Wales into factors underlying excess winter mortality required anonymised individual and household level data from the Valuation Office Agency (an agency of HM Revenue and Customs (HMRC)). The request was submitted in April 2010, and legal teams discussed appropriate ways of providing the data. In February 2012 HMRC confirmed that they believed they were unable to share the data due to statutory constraints on data sharing. This had a negative impact on the implementation of government policy that might have had an impact on mortality. HMRC was subsequently able to make this or similar information lawfully available to another government department where the disclosure fulfilled the current legal restriction imposed on HMRC for disclosure of data, that it served HMRC's statutory purposes.

- **Understanding trends in Housing Benefit spending**

The Single Housing Benefit Extract Data (SHBE) is an extremely rich dataset on receipt of housing benefit. The Department of Work and Pensions (DWP) recently commissioned a research group to carry out an evaluation of the housing benefit reforms using this data. They gave the research group permission to store the data in an Enclave where there have been no security breaches over a two year period. However DWP are not granting permission to the research group to use this data for any other purpose and no other external researcher is permitted to use the data either. There are simple pieces of analysis which could be done with this data, not done in DWP, which would help understand trends in Housing Benefit spending.

- **Understanding the productivity gap**

A research project has been proposed to consider UK productivity and growth and understand the productivity gap. To do this the research would like to link census data to understand people characteristics, HMRC and DWP data to understand where the individuals work, their earning and social welfare status, HMRC data on profitability of organisations, and DfE data on qualifications. Using de-identification and linkage methods, the researcher can look at the links between the education of the population and the profitability of firms and the links to productivity. This study, which sought to link data from the Census to administrative records, has not been possible. The experience of the researcher is that it will take years (if ever) before data would be made available for such a study through existing narrow gateways for sharing.

Many previous studies demonstrate the value that emerges when we link the rich and extensive information collected about individuals through their participation in social and biomedical surveys, with the detail that comes from administrative data held about the same people. Here are some examples:

- **Discovering the link between smoking, lung cancer and cardiovascular disease**

The most famous study is by [Doll and Hill \(1954\)](#), who surveyed doctors from the British Medical Association over time about their smoking habits and eventually linked this to their death registration data to obtain cause of death. This ground-breaking study was one of the first to show the link between smoking, lung cancer and cardiovascular disease.

- **Assessing the impact of streaming on primary school children**

[Campbell \(2014\)](#) used the Millennium Cohort Study and data from the National Pupil Database to assess the impact of streaming on primary school children. She found evidence to suggest that teachers' perceptions seem to be influenced by streaming in a way that advantages pupils in higher groups and penalises children in lower placements.

- **Understanding prevalence of self-harm**

Mars (2015) used the Avon Longitudinal Study of Parents and Children to investigate the levels of agreement between self-reported and medically reported self-harm in a cohort of adolescents, by linking survey data to administrative records held by the NHS. The research found evidence that self-harm prevalence may be underestimated if relying on self-reported data only.

Finally here are some examples of studies that the Administrative Data Research Network¹ is working to progress, although access to the data in some of these cases has constrained their progress and continues to do so.

- Evaluating linkage between children's health, education and social care data - pre-term births sub-study (<https://adn.ac.uk/research-projects/approved-projects/project015/>)
- An Investigation into the Impact of Disability on Employment in Wales (<https://adn.ac.uk/research-projects/approved-projects/project016/>)
- Evaluating the fairness of admissions decision making in UK higher education (<https://adn.ac.uk/research-projects/approved-projects/project107/>)
- Exploring associations between job characteristics and objective and subjective measures of well-being (<https://adn.ac.uk/research-projects/approved-projects/project113/>)
- Fluctuations in Community Mental Health and Social Care Services for people diagnosed with a mental illness in England (<https://adn.ac.uk/research-projects/approved-projects/project072/>)
- The wider impacts of benefit sanctions: educational attendance, behaviour and attainment -(<https://adn.ac.uk/research-projects/approved-projects/project052/>)

¹ <https://adn.ac.uk/> The Administrative Data Research Network is a UK-wide partnership between universities, government departments and agencies, national statistics authorities, the third sector, funders and researchers. The ADRN help [accredited researchers](#) carry out social and economic research using linked, de-identified [administrative data](#) – information which is routinely collected by government organisations.

References

- [1] Parliament UK (2016) 'Digital Economy Bill 2016-17' (webpage) Available at: <https://www.gov.uk/government/collections/digital-economy-bill-2016>
- [2] DCMS (2016) *Factsheet – Digital Government: Research and Statistics* (PDF) Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/535021/11_Research_and_Statistics.pdf
- [3] ESRC (2014) 'Public dialogues on using administrative data' (webpage), Available at: www.esrc.ac.uk/public-engagement/public-dialogues/public-dialogues-on-using-administrative-data/
- [4] 'RSS research finds data trust deficit ...' *StatsLife*, 22 July 2014. Available at: <https://www.statslife.org.uk/news/1672-new-rss-research-finds-data-trust-deficit-with-lessons-for-policymakers>
- [5] 'Public supports sharing health data for research...' *StatsLife*, 14 March 2016. Available at: <https://www.statslife.org.uk/news/2721-public-supports-sharing-health-data-for-research-poll-says>
- [6] 'New safeguards and public conversation about health and care data proposed', *Gov.UK*, 6 July 2016. Available at: <https://www.gov.uk/government/news/new-safeguards-and-public-conversation-about-health-and-care-data-proposed>
- [7] Professor Sir Charles Bean (2016) *Independent Review of UK Economic Statistics* (PDF) Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/507081/2904936_Bean_Review_Web_Accessible.pdf
- [8] 'UK official stats largely in compliance...' *StatsLife*, 30 March 2015, available at: <https://www.statslife.org.uk/news/2164-uk-official-stats-largely-in-compliance-with-eu-code-of-practice>
- [9] Institute for Fiscal Studies (2016) 'What and where you study matter for graduate earnings – but so does parents' income' Available at: <http://www.ifs.org.uk/publications/8235>
- [10] Eg see 'the proposed strands' in Open Rights Group (2016) 'Better Data in Government Consultation'. Available at <https://www.openrightsgroup.org/blog/2016/better-data-in-government-consultation>
- [11] Scottish Government (2012) *The Guiding Principles for Data Linkage*, available at: <http://www.gov.scot/Topics/Statistics/datalinkageframework/GuidingPrinciples>
- [12] Science & Technology Committee (Commons) (2016) 'Big data' is a big opportunity, but Government should tackle misuse' (webpage). Available at: <http://www.parliament.uk/business/committees/committees-a-z/commons-select/science-and-technology-committee/news-parliament-2015/big-data-dilemma-report-published-15-16/>
- [13] Law Commission (2014) *Data Sharing between Public Bodies: A Scoping Report* (PDF). Available from: http://www.lawcom.gov.uk/wp-content/uploads/2015/03/lc351_data-sharing.pdf

[14] Administrative Data Taskforce (2012) *Improving access for research and policy* (PDF). Available from: https://www.statisticsauthority.gov.uk/wp-content/uploads/2015/12/images-administrativedatataskforcereportdecember201_tcm97-43887.pdf

[15] BIS (2013) 'Administrative Data Taskforce report: government response' (webpage). Available at: <https://www.gov.uk/government/publications/administrative-data-taskforce-report-government-response>

[16] UK Statistics Authority (2016) *Delivering better statistics for better decisions: Why we need new legislation for better access to data* (PDF). Available from: <https://www.statisticsauthority.gov.uk/wp-content/uploads/2016/03/Delivering-better-statistics-for-better-decisions-data-access-legislation-March-2016.pdf>

[17] Royal Statistical Society (2016) *Final response to the Cabinet Office consultation on better use of data in government* (PDF), available from: <http://www.rss.org.uk/Images/PDF/influencing-change/2016/rss-response-cabinet-office-better-use-of-data-consultation-april-2016.pdf>

September 2016