

Digital Economy Bill: Written evidence submitted by CLOSER.

1. About CLOSER (Cohort & Longitudinal Studies Enhancement Resources)

1.1 The ESRC and MRC funded CLOSER Partnership is a collaboration of eight UK longitudinal studies, the British Library and the UK Data Service. CLOSER's mission is to maximise the use, value and impact of the UK's longitudinal studies to help to build a better picture of people's lives across generations.

1.2 The UK's longitudinal studies are recognised as one of the most crucial sources of evidence for policy development in health, education, economics and society as they provide insights about change that cannot be obtained from any other data sources.

1.3 There are currently eight studies in the CLOSER Partnership, with participants born as early as the 1930s to the present day:

- Hertfordshire Cohort Study
- 1946 MRC National Survey of Health and Development
- 1958 National Child Development Study
- 1970 British Cohort Study
- Avon Longitudinal Study of Parents and Children (Children of the 90s)
- Southampton Women's Survey
- Millennium Cohort Study (Child of the New Century)
- Understanding Society: The UK Household Longitudinal Study

2. Evidence summary

2.1 This submission is in response to the Digital Economy Bill 2016 Chapter 5: Sharing for Research Purposes, as this is related to our work and expertise. It highlights the value of linking survey information with government administrative data; explains the existing procedures and safeguards for storing and accessing data; recommends that health data is included within the Bill, and that stronger incentives are put in place to encourage government departments to share administrative data for longitudinal research purposes.

3. Data linkage

- 3.1 The UK longitudinal study community are keen to embrace the potential offered by survey and administrative data linkage but are being hindered by cultural barriers within government departments and disparate and opaque data access processes and procedures.
- 3.2 Considerable value 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. These routine records provide objectively measured information about individuals interactions with service providers (e.g. exam results, GP diagnoses), whereas survey information provide nuanced self-reported information explaining individuals' drivers, aspirations and behaviors.
- 3.3 Therefore, while longitudinal studies collect their own data directly from study participants, many often augment this with data from administrative records (including medical records). This linkage further enhances the value of these publicly-funded datasets and improves the potential for these datasets to inform public policy development. Linkage is also beneficial to those with an interest in administrative data as it allows their interpretation to be informed by the richness of survey data.
- 3.4 Longitudinal studies also invest considerable resource in stakeholder involvement; it is crucial for longitudinal studies to understand public expectations and to maintain participant trust. Studies therefore are able to use public feedback to steer the development of data security and safeguarding facilities that allow research on linked routine records while maintaining legal compliance and public acceptability.
- 3.5 Researchers are interested in administrative data because its volume and detail can vastly exceed what is possible to collect through other routes. Research using administrative records can be of wider public benefit – for example, work on understanding the causes and consequences of social change, or exploring the impact of policy and practice on individuals and their families. As a result, bodies like the Economic and Social Research Council have set up special centres such as the Administrative Data Research Network (www.adrn.ac.uk/) to help facilitate and promote its use.
- 3.6 CLOSER is helping to securely link data from longitudinal studies to other records, such as NHS Hospital Episode Statistics and the Department for Education's National Pupil Database, so we can better understand how different aspects of people's lives interrelate. Examples of ongoing work in this area include linking longitudinal data to administrative records held by HM Revenue and Customs and the Department for Education, and to records held by NHS Digital.

4 Existing safeguards for accessing data

- 4.1 Allowing access to personal administrative data understandably raises important issues about public understanding and acceptability, confidentiality and autonomy.
- 4.2 There are robust procedures for accessing the data from longitudinal research. The Economic and Social Research Council (ESRC) funded studies are licensed through the UK Data Service (UKDS) to bona fide researchers for not-for-profit use. De-identified data are deposited for use by the research community at the UK Data Archive, based at the University of Essex, and available via the UK Data Service. For more potentially sensitive or disclosive, but still anonymous, information the UKDS operates an application-based secure lab facility for accredited UK Higher Education Researchers only. The UK Data Service use the 5 safes framework, also used by the Office for National Statistics (ONS) and HM Revenue and Customs, allowing access to data whilst protecting the privacy of the individuals represented in the data (see <https://www.youtube.com/embed/MIn9T52mwj0>).
- 4.3 All studies funded by the Medical Research Council (MRC) must comply with the “MRC Policy and Guidance on Sharing of Research Data from Population and Patient Studies” (see <http://www.mrc.ac.uk/publications/browse/mrc-policy-and-guidance-on-sharing-of-research-data-from-population-and-patient-studies/>). This sets out the MRC’s requirements and expectations for studies on matters including data standards, data sharing, the governance of data access, facilitation, and data-sharing agreements.

5 Case studies

- 5.1 Seminal work carried out by Doll and Hill (1954) provides a good illustration of the benefits of linking longitudinal and administrative records. They surveyed doctors from the British Medical Association (BMA) over time to find out about their smoking habits and eventually linked these responses to death registration data showing cause of death. This ground-breaking study was one of the first to show the link between smoking, lung cancer and cardiovascular disease and ultimately was instrumental in shaping public health recommendations and saving people’s lives.
- 5.2 More recently, the then Department of Business Innovation and Skills (BIS) wanted to identify socio-demographic patterns of entry into Higher Education. To do this they firstly commissioned Crawford et al¹ to use administrative data collected by the Department for Education and the Higher Education Statistics Agency to describe patterns of entry differentiated by different social backgrounds. A strong finding of this research was that ethnic minorities, females and those from the highest socio-economic backgrounds are most likely to enter university, and conversely, white, males and those from the lowest socio-economic backgrounds are least likely to go to university. However, the administrative datasets being used do not collect the full range of information required to explore the factors underpinning these differences. To do this, BIS commissioned Bowes et al² to use individual administrative records linked to survey data collected as part of

¹ https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/474273/BIS-15-85-socio-economic-ethnic-and-gender-differences.pdf

² https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/474269/BIS-15-462-understanding-progression-into-higher-education-final.pdf

the Avon Longitudinal Study of Parents and Children and the Longitudinal Survey of Young People in England. This linkage yielded important insights, including the role of attitudes and aspirations in the decision-making process and the different attitudes that prevail in different disadvantaged groups.

6 Digital Economy Bill: Recommendations

6.1 We welcome the proposals for data access and research in Part 5 of the Bill and that linking survey and administrative records are enabled through this piece of legislation, but have a number of concerns with the Bill as it stands.

6.2 Health data: We remain concerned that the Bill as it stands may be interpreted as meaning health data is excluded and cannot be linked with other administrative records under the power of this Bill, where the conditions of the Act regulating the sharing of health records are also met. Data linkage across health and other domains underpins hugely valuable research of considerable public benefit and should not be prevented by departmental differences in data access which risk creating 'silos' containing different classes of information.

6.3 Incentivising data sharing: The Digital Economy Bill stops short of requiring government departments to share their administrative data. Our view is that more robust encouragement of sharing will be necessary before we get near to extracting the full potential from administrative data. We would like to see consideration of stronger incentives in the Bill to encourage departments to actively participate and to overcome aversion to low levels of residual risk.

24 October 2016