

1.1.1 RSS RESPONSE TO ONS CONSULTATION ON THE FUTURE OF POPULATION AND MIGRATION STATISTICS IN ENGLAND AND WALES

26 October 2023

2 Summary

2.1.1 The Royal Statistical Society is a professional society for statisticians and data scientists, with over 10,000 members. Our members have a diverse range of knowledge and experience and we have consulted with them to inform this response.

2.1.2 The RSS broadly supports the ONS's aim to replace the decennial census with annual population estimates based on admin data – there is a strong case for attempting to do this based both on the potential to provide more consistently accurate population data (as census data becomes increasingly inaccurate over the ten years between censuses) and on cost. We note that there have been some excellent developments in the use of administrative data for producing population statistics – the Dynamic Population Model, while something of a “black box”, appears to be able to produce better quality estimates using a wider range of sources than has been possible in the current system. If this type of success can be built on and administrative data can be used to more regularly provide reliable data without requiring censuses, then that would be a great achievement for the UK's statistical system.

2.1.3 There are six areas where we have concerns about this approach:

1.1.3.1 Data sharing: The approach is dependent on reliable data sharing across government. Has consideration been given to the possibility that some data stops being available for sharing? Or that data may change as a result of administrative changes?

1.1.3.2 Census-dependent statistics: Some other official statistics depend on census data – might users of these statistics be impacted by the planned changes?

1.1.3.3 Breadth of data: Will the administrative data cover all the same areas as census data, and in a holistic way?

- 1.1.3.4 Depth of data: Census data provides quality information at the sub-local authority level – does the ONS plan to continue to deliver this level of information and, if so, how?
 - 1.1.3.5 The role of surveys: There is an indication that surveys will be required in order to replicate some of the information from the census – how confident can we be that these will be able to provide the same level of detail as the census in a cost-effective manner?
 - 1.1.3.6 Coherent UK outputs: Given that Scotland’s and Northern Ireland’s statistical offices may not follow the approach of England and Wales, how would coherence of UK-wide population data be ensured?
- 2.1.4 None of these areas of concern are insuperable – but it is important for users of census data and the datasets that rely on census data to be confident that these areas have been properly considered before a final decision to stop the decennial census is taken. The following section provides further detail on the six areas where the RSS would like to see more detail of ONS’s plans.

3 Detailed comments on the ONS’s proposals

3.1 Data sharing

- 3.1.1 While the RSS is supportive of the ONS’s direction of travel, using administrative data to replace the census is reliant on data sharing across government and we welcome the extensive research that has been undertaken and published to support the proposed way forward. It will require robust access to and control over administrative data from government as well as allowing data to flow to other government departments and the devolved administrations. However, ONS’s current access to administrative data appears to be fragile and under resourced. There does not currently seem to be a system in place to address this – is the ONS confident that this situation will be resolved before 2031? (In fact, probably long before 2031 to be able to inform the decision.)
- 3.1.2 Assuming that this can be addressed, since current access to administrative data is tentative, there is also a question as to what happens if a data set becomes unavailable. What is the



backup, once in production, if some data becomes unavailable? Might there be changes to the frequency of outputs for example or a temporary drop in quality? There is also a risk that data may change fundamentally due to policy changes in a government department over which the ONS has no control. For example, what would happen if there were significant structural changes to the welfare system or the tax system that changed the frequency or granularity or even existence of a series that the ONS relied on? Using multiple sources of information can, perhaps, allow the new system to cope with the loss of one data source – but this system will have to serve us for many years and in twenty years or more the situation could look very different. Rather than ‘waiting and seeing’, which would accept that statistical use of administrative data is secondary, we suggest that ONS should find ways of registering the importance of statistical use and establishing a role in the future design and delivery of administrative data systems.

- 3.1.3 The only reference in the consultation document to this type of challenge to statistical integrity that we can see is a statement that “The ONS is maintaining an ongoing dialogue with data providers to understand their operational changes, and would continue to work in partnership with them to ensure the data gathered is as relevant and reliable a source for statistics as possible” (p.34). It might be worth looking at other jurisdictions to see how engagement in administrative data development is guaranteed. For example in Ireland, the [Statistics Act, 1993](#) Section 30, gives the Central Statistics Office (CSO) a right to access to records held by other authorities and Section 31 requires any authority changing their collection or systems to consult CSO. Changes can still be made, but the requirement to consult at least stops sudden catastrophic loss of admin data. There may well be other examples of legislation which hard wires consideration of the statistical use of administrative data in countries which are based on a register. This is a substantial risk and we think greater engagement is needed around it.

3.2 Census-dependent statistics

- 3.2.1 Census data will not be the only dataset affected by this change – there are other, important, official statistics that rely upon the census. It provides an independent benchmark for mid-year population estimates so an external benchmark would be needed for administrative-based estimates from the Dynamic Population Model – this is acknowledged in the consultation, but a benchmark is not identified. The census is being used as the key dataset to calibrate and



validate methods to produce data products from digital footprint data, for example in the £59 million [Smart Data Research UK programme](#). And another important example is the Index of Multiple Deprivation – this is a census-dependent dataset that is important both to the Department for Levelling Up, Housing and Communities – but also to the parts of wider civil society that are concerned with supporting the least well-off in society.

- 3.2.2 There is also an issue of rebasing. Currently population estimates are rebased every ten years using the census. There is always a gap between rolled forward population estimates and the census – reflecting inherent biases in both. These are often systematic and reflected in particular groups of the population, such as young male migrants. Under the current proposal the last census would be the one undertaken in a pandemic, during a period of lockdown reflecting a different pattern of "usual residence" than could be expected in the subsequent years.
- 3.2.3 It is important that the impact on users of census-dependent data is planned for so that any discontinuities in supply or changes to the data can be minimised and clearly communicated in advance and that any biases due to changes in rebasing can be accounted for.

3.3 Breadth of data

- 3.3.1 The census provides a uniquely broad snapshot of life in England and Wales – providing information on ethnicity, religion, occupation and qualifications that is otherwise very challenging to gather. The RSS understands that the ONS's ambition is to be able to replicate the breadth of data that is given by the censuses annually – we think that it is vitally important that this is achieved.
- 3.3.2 This is a significant challenge. There are no alternative sources for many of the self-reported variables from the census. Ethnicity is not routinely captured in other sources and, where it is, it is not done well or consistently. It is important that ethnicity should be used in combination with national identity, language and religion – as it is in the census – to gain adequate insight into communities. There are no other sources for information on: sexuality, gender, religion, unpaid care, occupation or household relationships. Self-reported health is not captured in other sources with full population coverage. All of these variables are essential for measuring/evaluating intersectionality, outcomes and impacts of policies.



- 3.3.3 Data linkage is mentioned many times in the consultation document. This is not surprising - one of the background papers to the document notes that “linkage of alternative data will inevitably play a foundational role in the transformed statistical system”. However, “the quality implications specific to linkage of administrative data are far less well-understood”, as acknowledged in the background paper. This is therefore another area in which we would expect to see wider consultation and engagement, including to explore public attitudes to ways of improving the quality of linked data in Statistical Population Datasets.
- 3.3.4 It is important that before any decision on the future of the census is made, thorough research is conducted on alternative approaches for capturing the type of information that we currently rely on the census for. We think it would be premature to make a final decision on the future of the census before a detailed proposal is put forward on capturing the breadth of data provided by the census. We would also like to see thorough engagement with users of this data to make sure that the new methodology will adequately meet their needs.

3.4 Depth of data

- 3.4.1 The consultation document states that the “ONS’s baseline objective for all topics currently covered by the census is to deliver annual statistics at the local authority level with a 12-month lag from the reference date” (p.24). Our concern is that this does not provide the depth of data that some users of the census need – those users who rely on the census for quality sub-local authority social, economic and housing statistics would not be well-served if only the baseline objective is met. This is important, in particular, for users in local government and civil society who need to have an understanding of differences within a local authority.
- 3.4.2 A key role of the census is to deliver high quality statistics for small areas within local authorities and the proposal guarantees this for only a subset of topics. Some important topics are missed out, for example, there is no guarantee of the production of statistics about housing tenure within neighbourhoods. This is a key statistical output by which local authorities understand and regulate the local housing market (and it supports research on housing welfare and suburbanisation of poverty).
- 3.4.3 It is important that an analysis of the uses and potential uses of sub-local authority data is made before any decision is taken to only provide given statistics at the local authority level.



3.5 The role of surveys

- 3.5.1 Delivering the full range of statistics currently provided by the Census will require an ongoing role for social surveys, but this role has not been specified. We note that research is still ongoing into a substantial number of areas – and that surveys will likely play a role in a significant number of cases. At this stage, the quality of the outputs cannot be predicted, and users cannot assess whether the proposed outputs could meet needs. It is important that the research is completed and reviewed with users before a final decision on the future of the census is made.
- 3.5.2 There are also questions around the modes of surveys, which will be an important consideration. For example, how many will need to be completed face-to-face or over telephone, versus by electronic questionnaires. How will they ensure against exclusion of the hard to reach and subpopulations such as speakers of other languages?
- 3.5.3 It is also important to recognise the wider context regarding surveys – where we are seeing a continuing declining in the number of people responding to ONS surveys. In the past month we have seen the Labour Force Survey delayed due to – presumably – quality issues arising from the low number of respondents. Labour market analysts and policymakers are aware that survey data and administrative data are giving different versions of the current tight state of the labour market. The problem is deepened because, currently, one of the tools that is used to weight survey responses when there is a low response rate is to weight results using population estimates from the census – without the census we will potentially be introducing additional uncertainty into the system. If surveys are going to be relied upon to provide information that was previously provided by the census, we would expect to see a compelling plan for increasing participation in ONS surveys and for ensuring representativeness. Similarly, it is not yet clear how coverage adjustment will be achieved for the admin-based population estimates (whose accuracy is heavily reliant on coverage adjustment). It is possible that this could be done using a large survey, but – if that is the approach – it is all the more important to address these concerns with the role of surveys.
- 3.5.4 This leads to a question about cost. The consultation document estimates that “the proposed transformation would cost less than half the cost of a 2031 census over a ten-year period” (p.40). It is unclear how this estimate is arrived at given both that the number of social surveys



required (and the effort required to induce people to complete them in sufficient numbers) remains unknown and that substantial effort may be required in order to encourage enough people to respond to them. It needs to be clear as to what the trade-off is between cost and quality. The savings may not be as great as anticipated, so it should be absolutely transparent before a final decision is made as to what would be saved in costs versus what would be lost in detail. It would be worrying if the estimate that the new system would cost less than half the cost of the 2031 census served as a budgetary cap – which may leave the new system underfunded.

3.6 Coherent UK-wide outputs

1.1.1 The consultation paper sets out a welcome commitment to “produce population estimates that are accessible, timely, and coherent, and which users understand, both now and in the future” (p.35) and we support that aim. However, there is currently not much detail on how that will be achieved should England and Wales move towards annual population statistics based on administrative data and surveys. This potentially poses a challenge that is much greater than that which is currently faced in producing coherent UK-wide population statistics. It would be helpful to understand how work in this area is being progressed, what mechanisms are in place to support it, and how users of UK-wide statistics are being consulted and engaged. Ultimately, how would the ONS manage should one of Northern Ireland or Scotland continue to use a traditional census?

4 Conclusion

4.1.1 The common thread across the six areas of concern that we have identified is that user engagement and clear communication are vital if the ambition of replacing the census with administrative data is to be achieved. The RSS would be happy to support this effort however possible and would welcome working with ONS and others to facilitate further and sustained engagement with users.

