Introduction to ONS projects supporting the National Statistician's recommendations on the future of the population statistics system

1 Introduction

In 2023, the National Statistician will deliver recommendations to Government on the future of the population, migration and social statistics system, including on the future of the census. This paper:

a. updates MARP on the structure of the projects supporting the delivery of these recommendations;

b. provides an update on the progress of work towards an integrated statistical design for the future population and social statistics system; and

c. provides an update on the status of the projects providing evidence for the recommendations.

It also contains a proposed forward look of items for MARP's consideration in the short and longer term (Annex B).

2 Structure of projects supporting the delivery of the 2023 recommendations

The work to deliver the 2023 recommendations on the future of the population, migration and social statistics system is broadly split into five areas:

a. Population and Migration Statistics Transformation (PMST), which is conducting research to produce admin-based population, migration and household estimates (ABPEs, ABMEs and ABHEs) using administrative data as a primary source, supplemented with survey data;

b. Social Statistics Admin First (SSAF), which is conducting research to produce outputs covering the full range of census 'characteristics' topics, using administrative data as a primary source, supplemented with survey data;

c. Requirements & Quality, which provides a cross-cutting support function to the above two projects, with respect to the acquisition of administrative sources and the assessment of their quality;

d. Social Survey Transformation (SST), which is producing a new social survey model to support PMST and SSAF's 'admin-first' approach to producing population, migration and social statistics;

e. The '2023 Recommendation Project', which aims to support the delivery of a set of recommendations based on the above work. The most important deliverables planned for this project include:

    i. An integrated statistical design for the future population and social statistics system;

    ii. Communications & engagement with users and the general public between now and December 2023, including a public consultation (see Annex A);

    iii. A set of policy considerations, including various impact assessments.

3   Update on progress of work towards an integrated statistical design

This section summarises the work that has been undertaken to explore the future population statistics system since the publication of the 2021 Census White Paper. A fuller background (from 2010-18) is provided at Annex C.

ONS' plans and research into the future of population statistics were included in Chapter 8 in the Census 2021 White Paper published in December 2018. This set out ONS' ambition to put administrative data at the heart of the population and migration statistics system. It also recognised the importance of looking at admin data for housing and household statistics.

Further it noted the potential of admin data to play a wider role in producing statistics and enabling analysis of topics not historically covered by the Census – "… An integrated data [administrative, commercial and survey data] approach offers the prospect of producing new, more responsive, timely and granular insights of our economy and society. This could include small-area household income distributions by health and ethnicity, fuel poverty, mental health, crime, inequalities, ageing, and housing affordability and provision."

ONS' plans for rebalancing its business and social surveys to support the administrative data-first approach were also described in the White Paper to:

- fill gaps that administrative data cannot fill. For example, some data are not available from administrative sources (e.g. hours of unpaid caring)
- help assess, and as necessary adjust for, definitional differences in administrative data where they don't directly meet what is required, particularly for key measures on the size of the population, migration and number of households
- reduce ONS' reliance on large population and business surveys, making remaining survey operations more efficient through a move from paper and personal interview-based collection to online

Recognising this, ONS moved away from the relatively narrow scope of the Admin Data Census (see Annex C) to look more holistically at the transformation of the wider population, migration

and social statistics system. The new project structure (set out in Section 2 above) was established in 2019.

The Population and Migration Statistics Transformation (PMST) project has continued to publish updates on ONS' ambition to put administrative data at the heart of estimating the size of the population and estimates of migration, the latest being in November 2020.

The project engaged with MARP to show progress on the high-level design and some of the research strands and dependencies.

- Integrated Population and Characteristics Survey – EAP119 here
- Estimating population size – EAP129 here, EAP 151 and EAP 159 (on Confluence, to be published soon)
- Statistical design – EAP132 here
- Fractional counting – EAP150 (on Confluence, to be published soon)

In addition, a similar programme of work is underway to transform the way we produce social statistics (the Social Statistics Admin First project, 'SSAF') to better meet the needs of our users and to produce the best statistics from all the available data. In the last year, we have published research on housing characteristics using Valuation Office Agency data, including how we can provide information on property type, property floor area and estimates of overcrowding.

4 Status report for research projects providing evidence for the 2023 recommendations

NB: dates mentioned in this section and in Annex B of this paper are provisional and liable to change

4.1 Population and Migration Statistics Transformation

We are continuing to develop our approach to using administrative data for population and migration statistics. This includes testing how our existing rules-based approaches work over time, and adapting these to work with new and changing data sources. We will be exploring approaches to producing population, internal migration and international migration statistics to ensure that any future system is coherent and consistent.

We are combining the insights we have gained from our work on admin data to date, and plan to explore how admin data have changed as a result of the pandemic to support delivery of the 2021 Census. We then intend to consider how we can incorporate new sources into existing admin-based population estimate (ABPE) methods. This will include exploration of individual linkage and approaches at aggregate level.

The introduction of the new online Labour Market Survey (LMS - formerly referred to at MARP as 'IPACS', and which includes population coverage survey questions) gives us the opportunity

to test the linkage of a large online survey to admin data. This will both inform our requirements for a survey in the future, and allow us to trial producing population estimates using a combination of admin and survey data. This will include testing the different flavours of dual system estimation approaches that we have previously shared with MARP, but will also include exploration of alternative approaches which deal with over-coverage (such as dependent interviewing). In addition, research has started to explore how to deal with the ongoing nature of the survey in the estimation framework, as well as expanding the simulation studies to provide additional realism.

We are also further developing our thinking around the framework that will be needed to assess the quality of statistics produced from an admin based system in future, building on the framework presented to the panel in EAP144.

4.2 Social Statistics Admin First

Within the Social Statistics Admin First project, research is currently being conducted under the following themes:
- Income and economic activity
- Education
- Housing
- Health
- Protected characteristics

This research aims to understand the quality and coverage of the admin data and develop methods to integrate admin and survey data to produce estimates. We are linking to existing survey and census data to develop these methods and will be learning from PMST's work to link the Labour Market Survey to admin data. We are also reviewing the requirements for social survey data collection on a wide range of population characteristics, particularly where admin data are not available or are of low quality.

Income and economic activity
Admin-based income statistics were [last published in December 2018](#) as experimental statistics. These covered employee and benefits income in England and Wales. The next development of these statistics, which will incorporate self-employment income, is due to be published in Summer 2021. These statistics use PMST's ABPE as their population base. Over the next year, we are planning to conduct research into imputation methods for producing income statistics covering the 2021 Census period. In conjunction with our admin-based income statistics, this will demonstrate the potential to produce small area income statistics that can be linked with data collected for the 2021 Census.

We are also planning to conduct research on admin-based labour market statistics over the next year, building on the early research that was published in [March 2018](#). We plan to publish our first developed admin-based labour market statistics in late 2021/early 2022, using the ABPE as the population base.

Education

Prior research, [published in October 2019](#), demonstrated that highest level of qualification could be derived for most new entrants to the labour market (as flagged in Section 3) from an experimental linked dataset created by the Department for Education, known as the Feasibility All Education Dataset for England (AEDE). For more information on the Feasibility AEDE, please see the following [publication](#). Current research is focusing on reproducing these results in ONS' updated computing environment, deriving FE/HE subject studied and level of apprenticeships (provisional MARP slot booked for June and provisional publication date in October), refining the methodology for deriving the highest level of qualification achieved and producing sub-national estimates. We are also working towards establishing an on-going supply of the AEDE and combining it with Welsh educational data sources to establish an All Education Framework.

Housing

We have previously published [research on housing characteristics using Valuation Office Agency data](#), including how we can provide information on [property type](#), [property floor area](#) and [estimates of overcrowding](#) (see Section 3). Notably we presented to MARP the Census 2021 edit & imputation approach for VOA number of rooms ([EAP120](#), [summary article](#) and [methodology working paper](#)) and how using VOA number of rooms (instead of asking a question) will affect the room standard, a measure of overcrowding and under-occupancy (EAP147, [publication](#)).

Current research is focused on:
- Developing and refining linkage methods for addresses and dwellings. Notably we are linking VOA data and Energy Performance Certificate (EPC) data to the Covid Infection Survey, as overcrowding is considered a driver of COVID-19 infections.
- Using EPC data to develop a statistical model for harmonising the VOA floor area measures for houses/bungalows and flats/maisonettes. If successful, future research will attempt to produce people per floor area as a new measure of overcrowding in England and Wales.
- Working with ONS' address index team to establish a residential address frame which could be used to produce experimental statistics on property age, property type, number of bedrooms, number of rooms and floor area using VOA data.
- Exploring the feasibility of linking VOA property type to the ONS Household Finance Surveys, to assess VOA quality and provide social surveys with property attribute information that is not currently collected.

- Working with LMS (part of SST) to explore if we could link VOA number of rooms to the LMS to derive the Bedroom Standard (a measure of overcrowding that uses the number of bedrooms in a dwelling and subtracts the number of bedrooms required, as determined by household information, to highlight if a dwelling is overcrowded or under-occupied).
- Acquiring and assessing new data sources to establish the feasibility of admin-based estimates of tenure.

Past research looked at how admin-data can be used to create small area estimates of fuel poverty. This work is currently on hold but may be picked up in the future.

Protected characteristics
Research is underway into the potential for producing estimates of the population by ethnic group from admin and survey data. We are planning to put out a publication in Summer 2021 using the 2016 ABPE as the population base and exploring the coverage and quality of Hospital Episode Statistics, Improving Access to Psychological Therapies and English School Census for providing data on ethnic group. Following that, further research will be conducted to explore different methods for treating ethnicity conflicts and refusals, expand on previous research using GSPREE to produce estimates of the population by ethnic group from admin and survey data, incorporate additional data sources for England, expand coverage to Wales (once additional data are acquired), and explore other methods as appropriate, such as imputation and fractional counting.

We are also looking to conduct research over the next year on other protected characteristics, such as religion, and other Census topics such as main language, car ownership and veterans.

Additionally, Methodology division are currently conducting research into fractional counting as a framework for producing small area multivariate statistics. A paper was presented to MARP previously on this (EAP150). Small area multivariate statistics are going to be a vital area for development, if we are going to meet the ambitious needs of our users, and provide statistics that give insights into policy outcomes.

4.3   Social Survey Transformation
An online only version of the Labour Market Survey (LMS) has been live since April 2020. The design is not finalised and is subject to change. Social Survey Transformation expect to engage with MARP in relation to the coverage elements of the LMS as and when the design evolves.

5   References

EAP104 – "Admin Data Census Population Coverage Survey", published working paper.

EAP111 – "Developing an ONS Population Spine", published working paper.

EAP119 – "Integrated Population and Characteristics Survey (IPACS)", published working paper.

EAP120 – "Imputing linked 2011 Census and administrative VOA data, A feasibility study", published working paper.

EAP129 – "Estimating population size without a census", published working paper.

EAP132 – "Statistical Design of Future Population and Social Statistics System", published working paper.

Butnaru, A., Gunner, L., Bull, S., Herson, J. and Tietz, S., (2021), "Admin-based levels of overcrowding (using the bedroom standard and Valuation Office Agency number of bedrooms) feasibility research: England & Wales: January 2021", published research.

Collyer, S., "Estimating the number of rooms in Census 2021: an update on imputation methods for Valuation Office Agency data", Published summary article.

Gunner, L., Dent, A. and Tietz, S., (2020), "Admin-based statistics for property type, feasibility research: England and Wales", published research.

Herson, J., Mason-Apps, E. and Tietz, S., (2021?), "Estimating the number of rooms in Census 2021: an update on deriving an occupancy rating from Valuation Office Agency number of rooms", published research.

HM Government, (2018), "Help Shape Our Future - The 2021 Census of Population and Housing in England and Wales", published 2021 Census White Paper.

Jones, P., (2019), "Research Outputs: Small area estimation of fuel poverty in England, 2013 to 2017", published research.

Letter from Sir Andrew Dilnott, (2014).

Mealor, A., Summerbell, A., Leather, F., Collyer, S., Dent, A. and Tietz, S., (2021?), "ONS working paper series no 20 – Feasibility of using donor-based imputation for census outputs on number of rooms using Valuation Office Agency data", published working paper.

ONS, (2014), "The Census and Future Provision of Population Statistics in England and Wales: Report on the Public Consultation", published report.

Skinner, C., Hollis, J. and Murphy, M., (2013), "Beyond 2011: Independent Review of Methodology", published report.

Taylor, B., Dent, A. and Tietz, S., (2020), "Admin-based statistics for property floor space, feasibility research: England and Wales", published research. Haines, N., Taylor, B. and Tietz, S., (2019), "Admin-based qualification statistics, feasibility research: England", published research

6 Annex A: Communications & engagement plans

The 2023 Recommendation communications and engagement plan intends to embrace a 'Test, Learn, Adapt' methodology to ensure our approach continues to deliver the best return on investment. It follows the Government Communications Service's OASIS model and draws upon Cabinet Office guidance.

The plan uses a phased approach to meet our aim to achieve broad acceptance of the recommendation and the benefits of a transformed system. We are committed to involving stakeholders in the development process throughout our path to the 2023 Recommendation, through discussion of in-progress work, opportunities to provide feedback and more intimate, themed engagement. This approach will ensure that we understand our stakeholders' views and concerns while also helping us to gather evidence to help shape our key decisions, and a basis to build support and advocacy for the transformed system.

7 Annex B: Forward look of items to be presented to MARP in 2021/2022

PMST	SSAF
During 2021/22 we intend to bring papers to MARP that cover the following: • Time series analysis of ABPEs from 2016 – 2019/2020 and what this means for their future development • Approaches to producing an admin based population proof of concept • A quality framework for assessing admin based population and migration statistics	During 2021 we have provisional plans to bring the following papers to MARP: • Admin-based Education Statistics, Apprenticeship Level & Subject Area • Admin-based Housing Statistics, EPC Floor Space • Admin data research on ethnicity • Admin-based income statistics • Transformation of housing questions for Household Surveys, feasibility of linking VOA property type to HFS • Modelling Health and Disability Status using admin Data SSAF research plans for 2022 are currently in development. Provisional dates will be shared once known.

8   Annex C: Background (from 2010-2018)

In 2010, the UK Statistics Authority asked the National Statistician and Office for National Statistics (ONS) to review the future provision of population statistics in England and Wales in order to inform the government and Parliament about the options for the next census. Over the next three years from 2011 to 2014, ONS undertook research into new ways of counting the population in England and Wales, reviewed practices in other countries, engaged with a wide range of users, completed a three-month public consultation and commissioned an independent review of methodology led by the late Professor Chris Skinner of the London School of Economics.

The programme published over 30 research reports here exploring ways to provide estimates on both the size of the population in England and Wales at small area levels and characteristics of that population. These included administrative data led approaches supported by new surveys, a 'traditional' decennial census and a rolling census.

The public consultation in 2013 covered two potential approaches for taking the census in the future. They were to take:

- a census once a decade, like that conducted in 2011, but primarily online

- a census using existing government data and compulsory annual surveys (a 4% annual survey was proposed)

Both approaches would have provided annual statistics about the size of the population, nationally and for local authorities. A census using existing data and surveys would provide more statistics about the characteristics of the population every year. An online census would provide more detailed statistics once a decade.

In 2013, the project commissioned Professor Chris Skinner to undertake a review of the methodology. This review made the following recommendations:

- there was a low risk to the online census approach, noting that it would represent a natural evolution of the traditional decennial census, paralleling developments elsewhere.
- there was a higher risk to the administrative data use approach, pointing out the problems of population estimation in a country without a population register, and the prior requirement for enhanced data-sharing between public bodies holding personal data and ONS.
- that further research be carried out into the administrative data use approach and identified the need for ONS to demonstrate that the proposed methodology could work. The review did not recommend substituting it for the decennial census yet.

ONS published a report on the consultation in March 2014. Whilst respondents varied in their views, there were some consistent key messages. In particular:

- there was continuing demand from government, local authorities, public bodies, business, the voluntary sector and individual citizens for detailed information about small areas and small populations offered by the decennial census, whether online or paper-based;
- there was a strong concern that the proposed use of an annual survey of 4 per cent of households (to support the use of existing administrative data) would not meet these needs, nor deliver the required small area and small population statistics offered by the decennial census;
- the more frequent statistics that could be provided between censuses by the use of administrative data and annual surveys would be welcomed, but not at the expense of the detailed statistics;
- whilst the methods using administrative data and surveys show considerable potential, there was concern that these were not yet mature enough to replace the decennial census; many respondents proposed a hybrid approach, making the best of both approaches, with an online census in 2021 enhanced by administrative data and household surveys.

As a result, in March 2014, the National Statistician recommended

- An online census of all households and communal establishments in England and Wales in 2021 as a modern successor to the traditional, paper-based decennial census. ONS recognises that special care would need to be taken to support those who are unable to complete the census online.
- Increased use of administrative data and surveys in order to enhance the statistics from the 2021 Census and improve annual statistics between censuses.

This was endorsed by the Chair of the UK Statistics Authority, Sir Andrew Dilnott, in his [letter of 27 March 2014](#) to the Minister for the Cabinet Office, Francis Maude. The Government responded on 18 July 2014, welcoming the recommendation but noting "Our ambition is that censuses after 2021 will be conducted using other sources of data... In the period up to 2021, the UK Statistics Authority's plans should include ensuring that adequate research into the use of administrative data and surveys is carried out to enable a decision about the future methodology for capturing population and census data."

The Census and Data Collection Transformation Programme (CDCTP) was established in 2015 to take forward the National Statistician's recommendation. The Administrative Data Census (ADC) project (part of CDCTP) was started to explicitly take forward the recommendations on administrative data. This continued the research on population estimation but was also expanded to include the potential of administrative data to provide information on characteristics of the population and to further explore the design of a population coverage survey to support population estimation from administrative data.

The project published a number of research reports and research outputs beginning in 2015 demonstrating the potential of a new administrative data census. These included:

- population estimates (for a range of geographies)
- internal migration estimates
- number and size of households
- household composition
- ethnicity (using combined survey and administrative data)
- income distributions down to small areas
- new mothers' income
- commuting flows using anonymised mobile phone data
- labour market status (including more detail on undergraduate students)
- highest level of qualification achieved

Annual assessments were also published summarising progress starting in 2016. The passing of the Digital Economy Act, giving ONS new data-sharing powers, was a notable milestone and met one of the findings from the earlier Beyond 2011 Consultation and Skinner Review.

The ADC project also engaged with MARP from 2018 onwards on some of its research:

- Population Coverage Survey – EAP104
- Developing an ONS population spine – EAP111