Quality focus for population and migration statistics transformation

Purpose

This paper aims to increase awareness of the quality research topics associated with the transformation of population and migration statistics, set out recommended work to address these gaps, and to gain this groups support for taking this work forward.

We group these research gaps into 3 key areas:

1. A robust strategy to improve the quality of administrative based estimates, requiring a coherent plan that unites the different strands of cross-office research more fundamentally.
2. Developing quality measures for the key inputs to and the outputs from the Dynamic Population Model (DPM), as well as methods for administrative data-based estimation. This is essential to ensure the quality is sufficient, allowing an informed switch between the mid-year estimates (MYE) produced by the current Census-based system to those produced by DPM.
3. Developing a quality framework and sustainable quality measures for administrative data-based characteristics estimates, progressing from case studies to more strategic and holistic research into the integration of administrative and survey data.

Recommendations

Members of the Methodological Assurance and Review Panel (MARP) are invited to:

- Provide assurance that:
    - A quality framework to deliver against the recommendations outlined to evaluate statistical quality in the Demographic Index and an increased quality focus on administrative data linkage is important.
    - Research and development of methods for estimation, accuracy and measures of uncertainty are fundamental for population and migration statistics transformation alignment to the Code of Practice and therefore are important next steps.
    - A quality framework for producing characteristics estimates via administrative and survey data is needed for a holistic development plan in this area. This could be facilitated by a meta-review of existing case studies, that could possibly identify a need for further case studies or generate generalised rules from which to develop a framework. We propose this meta-review and framework be delivered over/before further case studies into discrete variables.

Background

ONS' work to date using administrative data to produce population statistics has been successful and innovative. With a view to fulfilling the requirements of the National Statisticians 2023 Recommendation to government on the future of the Census for England and Wales (2023 Recommendation), the Data Architecture (DA) division has created the Reference Data Management Framework (RDMF), with indexes facilitating person matching (Demographic Index). The DI forms the essential spine of the Statistical Population Datasets (SPD); with SPD V4.2 currently in preparation for publication by teams in Social Statistics Transformation and Research (SSTAR). ONS has ongoing research into the incorporation of administrative data for estimating population characteristics or attributes. At the time of writing, experimental administrative data-based estimates for ethnicity, income, and

household variables have all been produced. For international migration statistics, the trajectory has been similarly ambitious. With the suspension of the International Passenger Survey (IPS) in March 2020, Migration Statistics Division (MSD) began a research programme into the use of the Registration and Population Interaction Database (RAPID) from the Department for Work and Pensions (DWP). Recent work has extended this to incorporate additional administrative data from the Home Office (visa and border data), some IPS data, and begun to rely less on statistical modelling. Finally, a Bayesian demographic accounting framework, referred to as the Dynamic Population Model (DPM), is under development within SSTAR and Methods and Quality Directorate (MQD) as an experimental alternative to existing mid-year population estimates (MYE). The DPM incorporates the SPD as the core data input for population stocks and brings together measures of natural population change (births and deaths) and movements (international and internal migration; see Figure A1 in the Annex). The work is currently at 'proof of concept' stage.

These components of the transformation of population and migration statistics[1] form the core evidence for the 2023 Recommendation:

1. The most recent version of the SPDs, including a comparison with Census 2021
2. Proofs of concept demonstrating the ability to produce more timely and frequent population and migration estimates
    - Demographic accounting
    - International migration
3. Proofs of concept demonstrating ability to produce estimates for a selection of topics
    - Housing
    - Income
    - Ethnicity
    - Multivariate combinations of the above
4. Proof of concept demonstrating ability to produce a longitudinal census cohort study (Census Data Asset)[2]
    - Refugee Integration Outcomes (RIO) study

These external components have been reviewed individually by MARP. In addition, previous external reviews are useful as a guide to what information users and stakeholders might expect to receive in a holistic consultation around the transformation of population and migration statistics such as is planned with the 2023 Recommendation Consultation. The most significant and recent external review is the Skinner Review report (October 2013). Although it would not be sensible, given the transformation changes that have occurred since it was written, to consider the Skinner Review as a prescriptive document that must be adhered to, it is still a considered, thoughtful, and useful reference point by which to sense check our current trajectory. It's important to note here that MARP has not been presented with a holistic view of the entire transformation to the same scale as the Skinner Review.

With a focus on the accuracy component of quality, the Skinner Review recommended several elements of future research as essential work (paraphrased):

[1] It is important to note that all ONS publications/updates on the transformation programme make explicit use of the need to incorporate additional data collection, or surveys, to compliment the administrative data.

[2] As the CDA is still at design development stage, it is not covered within this paper.

1. Quantification of the level of accuracy (bias, variance) associated with the administrative data option.
2. A statement on the nature of the likely estimation errors
3. Justification that the level of accuracy is appropriate; assessment of the likely accuracy of population estimation by age and sex and for areas smaller than local authority (LA); demonstrating that quality standards for population estimates at the national level are likely to be met.
4. An estimation methodology which realistically captures the effect of both
    a. Over-coverage in the SPD
    b. Matching error
5. Use of the estimation methodology to demonstrate that neither 4a nor 4b represents a serious threat to the population estimates meeting the quality standards

Table A1 in the Annex considers the recommendations of the Skinner Review against the components of the transformed system that have developed since the Review's recommendations were published. This demonstrates where we now have evidence gaps that require workplans to deliver against. Approximate time scales for delivering against these gaps are provided.

Discussion

The evidence for the 2023 Recommendation is made up of proof of concepts (excluding the SPDs which are far more established). If we consider the 'concept' of using administrative data to produce outputs for population and migration statistics to have been demonstrated, this paper outlines that to prove these concepts, we will now focus on producing measures of precision to understand whether their quality is suitable for users. ONS must be able to describe the impacts of quality on the use of the data and the outputs – this is not currently possible. For example, if an estimate of international migration is produced without sufficient information on the quality of that estimate, in terms of accuracy, or timeliness, etc, the risk is the misuse of that estimate in the public and policy domain with life-changing impacts for the population. Of immediate focus therefore must be research into accuracy across the board and a framework for developing characteristics estimates.

Accuracy

Prioritising the development of methods to assess accuracy and communicate uncertainty (variance, bias) against the gaps identified in Table A1 is needed. This will collectively be a large-scale, complex, piece of work. Without this, we will not get the full benefits possible from the DPM. One of the key differences between the DPM and current MYE system is that the former makes use of measures of uncertainty around its data inputs to decide which data source is more robust and therefore deserves more weight when reconciled against different estimates. Developing methods for measures of uncertainty is essential to maximise the nuances of the DPM.

The focus on accuracy will be established from the outset of the statistical lifecycle for the transformed system. MQD, in a paper shared today with MARP, has proposed several areas in which further work is needed: improving linkage methods for administrative data; improving data quality information including precision and recall metrics, for administrative linkages; improving the inclusivity of linkage; and development of a QA/validation approach for joins made via the RDMF.

As well as this, consideration will be given to how the errors feed through from data compilation to the statistical outputs that are made. Accuracy in administrative data doesn't only include dealing with linkage error, but also includes contending with data error and

coverage error. These aspects must also be addressed to avoid inaccurate outputs – we will invest in improving the quality of our main data sources to facilitate high quality linkage. Working alongside colleagues in the DI team, the Strategic Statistical Design team in MQD have proposed a large-scale piece of work to produce quality metrics on the DI. This paper has been reviewed by MARP who endorsed the suggestions. Recommendations here include establishing quantitative measures for quality in the DI; testing the DI design; establishing governance; and supporting users to use the DI appropriately. This refers to both external users of the DPM, and the ONS internal use of DPM population estimates in e.g., benchmarking of social surveys.

The next step is developing methods to produce measures of uncertainty against all the different components identified in Table A1 and making sensible and informed comparisons against the established accuracy standards. Of particular importance here are uncertainty measures for the inputs to the DPM (SPD, internal and international migration, the single sources).

We make specific reference here to improving the methods for SPD uncertainty measures (with/without Census 2021 as a reference point). We are clear here that there will be a need for an additional coverage survey to produce improved SPD estimates within an administrative data and survey estimation framework. MQD work in this area is currently focused on a rudimentary estimation methodology using the Census Coverage Survey (CCS), to deliver against the customer requirement to implement an SPD coverage estimation method by April 2023 in time for DPM outputs in June 23. This however will not provide an accurate indication of the quality achievable with the best possible method or with data from the Labour Market Survey (LMS), nor will it allow us to continue our research on a longer term, robust estimation strategy as preferred by MARP. This increases the importance of following an estimation framework.

Uncertainty measures for administrative data-based migration estimates (ABMEs) are especially important, given the highly experimental nature of these estimates and the high public and media interest. MQD research in this area has been commissioned by the business area but is still at the early stages of thinking. This work will be presented at MARP later in the Spring.

The uncertainty measures for the DPM outputs themselves must also be developed – currently we lack the ability to produce Bayesian credible intervals at the aggregated level which prevents a complete comparison with estimates from the MYE system.

To inform this DPM/MYE comparison, initial work within MQD has produced accuracy standards for the 2023 Recommendation in relation to Census 2021 and MYE 2020 presented to MARP today. The project initially focused on producing different standards for variance and to a lesser extent bias, at Local Authority (LA) level, for population estimates. No work has yet been undertaken to produce accuracy standards for the administrative data-based characteristics estimates to be assessed against. This work will be more complex and exploratory than producing accuracy standards for the population level estimates.

Finally, as mentioned, there remains a key gap in producing uncertainty measures for administrative-based characteristics estimates. This needs to be prioritised. Only then can informed decisions about the need for additional data collection for key variables be sensibly made.

Overall, there is a gap here for a strategic quality framework to shape the transformation and, also, an estimation framework to appropriately compliment the administrative data-based statistics. An increased focus on accuracy will allow more informed decisions on

potential trade-offs between precision, granularity, and timeliness and frequency, as well as inform the outputs of, and decisions following, a dual run between the DPM and current system for MYE[3]**.** A suggested framework for this could be the Quality Questions and Red Flags document from MQD Data Quality Hub and Analysis Standards and Pipelines Hub.

Additionally, we need to consider how to judge the extent to which changes in estimates between Census 21 and the administrative data option reflect artefacts of changes in methodology rather than genuine changes and, more generally, how to validate estimates. Coverage error, linkage error, and data error can all be conflated with genuine change. More broadly, a strategy is needed for if there is evidence that some of the estimates are of unacceptable quality. The validation exercise should be sufficiently substantial that it provides a means of producing alternative estimates if the population estimates produced by the system are deemed to be of unacceptable quality.

All the above listed quality research, including an appropriate estimation framework, must therefore be prioritised for DPM outputs to move from proof of concept, to experimental. Once all quality standards are available, an iterative approach to the transformation of population and migration statistics could be to identify a sensible minimum accuracy target, with iterative improvements, and embed these within the transition states. A dual run between the DPM and the MYE system can then inform whether the DPM outputs can move from experimental to being submitted for national statistic accreditation.

We therefore ask for agreement that:

- The work outlined to develop evaluating the quality of the DI and investing more research into the quality of admin data linkage is important
- Research and development of methods for estimation, accuracy and measures of uncertainty which are fundamental for population and migration statistics transformation alignment to the Code of Practice are important

Without robust measures of the accuracy component of quality, we cannot objectively say whether the transformed statistics are ‘better’ than the existing ones or provide users with evidence on any potential trade-offs between precision and frequency.

In terms of communicating uncertainty, there are several resources here whose guidance should be taken on. Guidance on communicating quality, uncertainty and change produced by MQD Data Quality Hub gives a good overview on this topic but references to administrative data are limited. A more recent release from the OSR on approaches to communicating uncertainty in the statistical system gives more robust advice specific to administrative data. The OSR also have standard settings documents and a QA toolkit for administrative data. Consideration needs to be given on how errors may propagate through the process of creating administrative based outputs and how to assess that cumulative impact on quality. Overall, we suggest that communicating accuracy, uncertainty and change to analysts and users from a perspective of how to handle and interpret quality measures for administrative data-based estimates should be developed by ONS. Again, this is important for both internal and external users of population estimates.

[3] Both the Office for Statistics Regulation (OSR) and Chair of the United Kingdom Statistics Authority (UKSA), Sir Robert Chote, have supported a parallel run of current MYE systems and the DPM before a move towards ‘switching off’ the MYE system.

Comparability

For statistics on population characteristics from administrative data, ONS teams in SSAF supported by MQD have produced univariate estimates on ethnicity, income, and housing, with some multivariate case studies in development. Key to these multivariate case studies has been the creation of linked administrative datasets, including the ABIES (Administrative based income and ethnicity statistics). The case studies and linked datasets have been developed specifically for the 2023 Recommendation. This work is novel and has been delivered against extremely challenging deadlines for which the teams must be commended. At the same time, the business area has expressed a concern regarding the quality of the linkage here, which again finds support in the proposal for a quality framework for administrative data linkage.

Feedback from MARP on reviewing the univariate and multivariate research has been that a holistic framework for producing population characteristics from administrative data should be developed, which includes the use of all available data sources including Census and social surveys. This integrated methods framework would allow comparison of methods and identification of cases in which additional data collection via surveys is required, rather than developing bespoke solutions based on individual variables.

This framework could be facilitated by a meta-review of the existing case studies. The review could identify a need for either further specific case studies or generate generalised rules from which to develop a framework. We propose that this meta-review and the development of a framework be developed prior to any further case studies into discrete variables.

We therefore recommend that, as well as the accuracy considerations for characteristics estimates listed in the above sections (i.e., essential research into sustainable quality measures), that:

- A framework for producing characteristics estimates via administrative and survey data be developed to allow holistic and coherent developments.

Conclusion

The paper has set out the main areas for a focus on quality around the transformation of population and migration statistics:

- A quality framework for the use of administrative data in population and migration statistics including advancing research on measures of uncertainty. This overlaps with other work proposing an increased focus on the quality of administrative data linkage and the DI.
- An estimation framework for the SPD which allows the DPM to move from proof of concept, to experimental, to final, estimates in a planned, quality-driven manner.
- A strategic, quality focused trajectory for population characteristics statistics planning for use of both administrative and survey data.

Katie O'Farrell (MQD)

Table A1. Skinner Review (2013) accuracy recommendations compared to current ONS workplans

Skinner Review recommendation	**ONS project currently delivering against this**	**Ready for inclusion in 2023 Recommendation User Consultation**	**Approximate size of work needed**
Quantification of the level of accuracy (variance and bias)			
In the DI	DI QA project from MQD	No	12 months: 2 months to get initial measures for whole-DI quality. Over 6 months, expect to begin: clerical review as standard to assess DI quality (including prevalence of errors, and uncertainty for these estimates) and simulation work to assess impact of DI error on analyses that use DI - simple cases. Over next year, expect to develop of quality measures for users - a first iteration, including what they mean for analysis
In the SPD	CAPS projects from MQD for variance SPD releases from SSTAR describe levels of over and under coverage against each release	No Yes	4 months: 2 months to generate SPD (4.2) uncertainty intervals for 2021, and further 2months to generate 2022 uncertainty intervals. This is using current method, w/o census estimation adjustment. N/A
In the characteristics linked datasets	None as yet	No	Not scoped out as yet
In the population flow data: births and deaths	Births and deaths counts do not incorporate uncertainty as they are assumed to be accurate, but when included in DPM as rates incorporate uncertainty from SPDv3 denominator. Future work may involve use of Census 21	N/A	N/A

In the population flow data: international and internal migration	ABME uncertainty project within MQD for international migration SSTAR working on internal migration but uncertainty not scoped into this work	Methods options paper due at MARP in January 2023 but won't be implemented No	For international migration: 4months to test after feedback and recommendation from MaRAG on the options paper Not scoped out as of yet
Additional single source inputs to DPM	Qualitative work from Research Hub in MQD (Inclusivity research for PDS and HMRC/DWP, Electoral Statistics quality review for PMST Lots of measurement errors, missingness etc etc) Future quantitative work on PDS from Research Hub in MQD Work from Quality and Requirements team in SSTAR – qualitative work and longitudinal change on PDS, including missingness and changes in supply	Potentially some papers from SSTAR but not definitive	Ongoing/Not scoped out as of yet
In the DPM outputs	From the DPM project team in MQD: At disaggregated level (LA	Yes for disaggregated level	Work in the Research Methods hub in MQD could produce approximate uncertainty measures for aggregate level outputs by June 23.

	by SYOA by sex) but not at aggregated level		
In the characteristics outputs	None as of yet	No	Not scoped out as of yet
Statement on the nature of the likely estimation errors	If we define this as relating to DPM outputs – none as of yet	No	Not scoped out as of yet
Estimation methodology which realistically captures the effect of both			
Over-coverage in the SPD	SPD Coverage project within MQD	Methods options paper passed by MARP Nov 22 for inclusion in 23 Recommendation consultation document	Implementation not scoped out as of yet
Matching error	None as of yet	No	Not scoped out as of yet
Use of the estimation methodology to demonstrate that neither over-coverage nor matching error represents a serious threat to the population estimates meeting the quality standards	SPD Coverage project within MQD	No	Not scoped out as of yet