

Developments in official statistics 1

Counting shadows: A community-anchored modelling framework for estimating unauthorised population size

Maciej Beręsewicz - Poznan University of Economics and Business; Statistical Office in Poznań, Poland, **Brendan Georgeson** - UK Office for National Statistics, **Peter Walsh** - Migration Observatory, University of Oxford, UK, **Aniela Czerniawska** - Poznan University of Economics and Business; Statistical Office in Poznań, Poland, **Li-Chun Zhang**, University of Southampton, Southampton, UK; **Statistisk sentralbyrå** (Statistics Norway), Oslo, Norway

In this paper, we propose a new method for estimating the size of unauthorised populations in a given country. Our approach is based on a theoretical and empirical model that uses only aggregated data on the observed unauthorised and lawful populations. The model assumes that the unauthorised population is a random variable conditional on the reference (authorised) population and is related to its size through social networks. Our approach is valuable as an estimation framework when data are limited or when access to individual-level data is not possible. We apply this approach to Poland from 2019 to 2024 using three reference populations. We estimate that the point estimate of the size of the unauthorised population in Poland was between 30,000 and 150,000, varying by distributional assumptions, and that its relative size to the reference population was between 5% and 20%. Beyond the Polish case, the approach offers a portable/transferable estimation framework for national statistical offices seeking to quantify unauthorised populations using limited, aggregated data. For the UK the data sources would be the Mid-Year Estimates (MYE), the recorded arrest data by nationality including Border Force data and Home Office data on voluntary and involuntary returns. The potential application would be to estimate the UK unauthorised population stock for inclusion in the MYE and to support estimates of flows of unauthorised migration by inclusion in the Long Term International Migration estimation. The underlying assumption of the model is 'ethnic economies' hence the title Community-Anchored Modelling.

Developments in the Office for National Statistics Longitudinal Study (ONS-LS)

Merilynn Pratt & Erin Lumsdale - Office for National Statistics

The ONS Longitudinal Study (LS), established in 1974, combines Census and linked life events data for a 1% sample of the population of England and Wales. We have published the Longitudinal Study (ONS-LS) Series 12 report providing detailed quality information on the recent ONS-LS Census 2021 linkage. From 2021, we lost our external tracing system, the Medical Research Information Service Integrated Database Administrative System (MIDAS), which supplied minor events data to the LS. In response we have started an ambitious development programme which includes the development of interim methods for several of the ONS-LS tables.

In this presentation we will give an overview of:

- the linkage rates, tracing rates and sampling fractions for the ONS-LS Census 2021 linkage
- outline our interim methods for the minor events tables (Immigration, Embarkation and Re-entry tables) using Personal Demographics Service (PDS) data
- outline our planned development work for the LS which includes;
- Scoping new data sources to improve or provide alternative methods for existing tables.
- Updating tables which have been paused for some time such as cancer registrations and enlistments.
- Reviewing systems and quality assurance measures
- Planning for the link to Census 2031

Beyond household projections: Developing families projections for England

Sam Mold & Elliott Larkinson - Department for Work and Pensions

The Department for Work and Pensions has been collaborating with the Office for National Statistics to develop bespoke projections of families for use in Family Resources Survey grossing and benefit forecasts. In this presentation, we talk through how this was a lot harder to do than initially expected, how we developed the projections and how we are using them.

Demographic projections for London's new Spatial Development Strategy

Ben Corr, Donal Ring, Sebastian Heslin-Rees, Izabel Bahia - Greater London Authority

The London Plan is an integrated framework for the development of London for the next 25 years. A new draft Plan is set for publication in June 2026.

Here we present an overview of the small-area population projections that form part of the evidence base for the Plan:

- * Models and assumptions: reconciling population trends with housing capacity; implications of changes to the NPPF

- * Preparation of model inputs: future housing capacity, local (synthetic) migration estimates, the new Housing Market Area geography, Age Specific Fertility Rates

Recent changes to the Standard Method for the calculating housing need specified in the National Planning Policy Framework (NPPF) complicate the creation of these projections. The new method diverges from previous guidance by making the calculation of housing need independent of demographic projections.

The previous consistency between (trend-based) projections of population growth and planned housing development had allowed several simplifying assumptions:

- * The overall population of London could be projected using a trend-based approach, with no need to explicitly account for additionality from new housing

- * Future housing delivery in London would only affect the distribution of population within the city

- * Housing delivery in neighbouring regions could be ignored

These assumptions are no longer defensible with the new method expected to result in housing targets for many areas that greatly exceed rates of population or household growth.

The GLA has modified its approach, expanding the area covered by the model beyond London and explicitly accounting for housing development across the wider Housing Market Area.

The challenges and possibilities around creating a housing-adjusted population projections tool

Elzemies Scott-Kortlever, Ash Riley, Jack McQueen - Office for National Statistics

When the 2022-based subnational population projections (SNPPs) were published in June 2025, some stakeholders asked us what the impact of their future plans on housing developments would be on their projections. Unfortunately, there is no England-wide dataset on future housing plans that can be used as part of the SNPP production system.

When in place, a housing-adjusted projections tool would enable users to input their future plans for housing over a period of time. This information would then allow the tool to adjust the SNPPs up or down depending on the input and a set of previously defined parameters. The project is a collaborative effort between several ONS teams and NHS and local government stakeholders.

We are using data from the 2011 and 2021 Census on people and households, as well as data on household projections, housing prices and affordability, mid-year estimates, data on vacancies and second homes and annual housing completions.

The methodology behind the tool requires us to overcome several challenges around data availability and limited timeseries. Our talk will focus on the issues the work has to overcome, the assumptions that we're having to make, how we are trying to make sure the tool works for a wide range of local authorities and what initial results are showing us.

Developments in official statistics 2

Regulating a Census: a reflection and forward look to OSR's regulatory approach for 2031

Adele Nethercott - Office for Statistics Regulation

Official statistics should serve the public good. The Office for Statistics Regulation (OSR), as the independent regulatory arm of the UK Statistics Authority, works to uphold the trustworthiness, quality and value of statistics and data used as evidence; protect the role of statistics in public debate; and develop a better understanding of the public good of statistics. Our regulatory work is guided by the statutory and recently refreshed Code of Practice for Statistics, version 3.0.

The census is one of the most important sources of data and statistics, informing decisions about almost every aspect of life within the UK. It gives users access to important information on the people and households of the UK and helps us better understand the places in which we live and work. Plans across each of the UK Census Offices for a Census in 2031 are currently underway. In this presentation we will reflect on our regulatory approach of the 2021 Censuses in Northern Ireland and England & Wales, and the 2022 Census in Scotland and share our lessons learned - and what this means for our approach to assess the Census 2031 outputs. Our role is to ensure that the statistics meet the standards of the Code, so that they remain trusted, are of high quality and valued by users.

Causes of fluctuations in Scotland's Administrative Data Based Population Estimates

David Rowley - National Records of Scotland

The Administrative Data Based Population Estimates (ABPEs) are important for supporting Scotland's Census in 2031. They closely match the official mid-year estimates and their distributions by age, sex and geographic areas. However, the ABPEs fluctuate more between years than the official estimates.

The ABPE's individual data sources were investigated at aggregate and individual level, in isolation and in combination with the other sources, to understand these fluctuations, and if and how they can be avoided to produce a better quality assurance comparator for Census 2031.

Various factors were found to cause the fluctuations. Variation in maintenance done on source datasets (e.g. the NHS focused list cleaning exercises) can affect how many people appear on the source dataset. Source datasets also vary in size as the source database's relevance varies (e.g. the electoral register typically increases around major elections). If the use of the corresponding service varies this can also affect the source dataset. For example, during lockdown fewer people interacted with the NHS, so the health activity datasets reduced in size). Finally, data-quality issues can affect particular sources, or how well they link with others, for particular reference dates.

Extra independent sources can mitigate these problems; had fewer sources been used, the ABPEs would have fluctuated more. Source datasets are quality assured in isolation on receipt. Some data-quality issues are only apparent after linking, so QA of sources should continue through the process. This analysis also identified further improvements to the ABPE process, including making use of address histories.

Geography assignment at record level on the Statistical Population Dataset

Marcus O'Brien, Rossella Icardi, Lydia Smith - ONS

The Statistical Population Dataset (SPD) integrates multiple administrative data sources to produce estimates of the usually resident population. Because different administrative sources can assign conflicting local authority (LA) geographies to the same individual, the SPD resolves discrepancies using a fixed source hierarchy. This research examines whether alternative hierarchies can improve the accuracy of geography assignment, whether non active Personal Demographic Service (PDS) records should be incorporated, and whether Dempster-Shafer Theory (DST) can provide a robust quality measure and support geography assignment in years where Census data are unavailable.

The study evaluates multiple hierarchies using SPD 2021 data, benchmarking performance against linked Census 2021 geography. Ten alternative hierarchies are tested, and DST is applied to generate record level quality scores and to assess its suitability for ongoing quality monitoring and geography assignment. Data

sources include PDS, School Census, Higher Education Statistics Agency (HESA), Department for Work and Pensions datasets, and other administrative sources used within the SPD.

Preliminary results indicate that the scope for improvement through hierarchy changes alone is limited: only 6.7% of SPD 2021 records contain conflicting LA information across sources. The existing hierarchy performs well overall, with most alternatives reducing Census match rates. However, targeted adjustments improve outcomes for specific age groups, notably children aged 2–17 and young adults aged 19–24. Overall, DST demonstrates clear value as a scalable, record level quality indicator and offers a practical approach for assessing SPD geography quality beyond Census years.

Using alternative data to transform travel and tourism estimates

Jake Argent – Office for National Statistics

Following a review of our statistics in 2019, the UK Office for National Statistics (ONS) are transforming the way we collect travel and tourism data to deliver more efficient, accurate and coherent estimates. The Travel and Tourism Reform Project is using administrative data to gain insight into travel and tourism inbound to and outbound from, the UK.

ONS is using linked borders and immigration data from the Home Office as well as financial transaction data, to better understand international mobility, transportation and patterns of spending inside and out of the UK. In collaboration with the University of Southampton, we are also developing a statistical model known as the Multidimensional Structure Preserving Estimator (MSPREE) to predict variables not available in administrative data.

The project aims to provide:

- robust quality assurance to existing survey-based estimates
- alternative indicators of travel and tourism
- increased resilience to the statistical system

To date, our research has shown that there is substantial promise, and benefits, in supplementing traditional survey methods with administrative sources; this includes work conducted to match variables, compare data and produce travel and tourism statistics.

This talk will cover our research to date as well as outlining how we plan to take this work forward in the future.