

# Longitudinal census data – research using the newly linked 2021/2022 Census data 1

---

**Tuesday 2 September, 4.45pm**

**How far do different Census output geographies yield different statistical results?**

**Ian Shuttleworth - Queen's University Belfast, Estelle Lowry - Queen's University Belfast, John Hughes - Northern Ireland Statistics and Research Agency, Neil Rowland - Queen's University Belfast, Aideen Maguire, Queen's University Belfast  
Sarah McKenna, Queen's University Belfast**

Output geographies change between Censuses to reflect spatial shifts in population and political/administrative redistricting. There is always a tension between revising geographical units to respond to new circumstances and keeping them the same to permit consistent analyses through time. This tension centres on the Modifiable Areal Unit Problem (MAUP); cutting up a population into units of different shapes and sizes can potentially yield inconsistent statistical outputs and misleading estimates of change. In some specialised contexts, like the Northern Ireland Longitudinal Study (NILS), temporal change is primary but even more standard questions about whether Northern Ireland (NI) has become more/less segregated or more/less uneven on other dimensions require cross-Census comparisons. Therefore, does the move to Super Data Zones (SDZ) in 2021 from Super Output Areas (SOAs) in 2001 and 2011 substantively matter? We answer this question by placing Census 2021 100m Grid-Square data on the SDZ and SOA geographies to calculate similar indices of segregation and to examine bivariate relationships. Furthermore, we assess, using the NILS, the impact of using SDZs rather than SOAs as proxy neighbourhoods as Level 2 in a ML model. We conclude that it is safe to make NI-level comparisons between Census 2021 using data counts for SDZs and Census 2011 using those for SOAs and that the transition to SDZs does not lead to a discontinuity in the NILS.

Email: e.lowry@qub.ac.uk

**Who Cares? The Social Profile of Unpaid Carers Over Time and Place**

**Maria Petrillo - University of Sheffield, Jingwen Zhang - University of Manchester, Matt Bennett - University of Birmingham, Gwilym Pryce - University of Sheffield**

This paper shares findings and learnings as part of ONS Census Longitudinal Study beta testing. Over 6 million unpaid carers provide the majority of support for people who have challenges related to long-term illness, disability or older age. The care that they provide is estimated to be worth £182 billion across the 4 nations of the UK - equivalent to a second NHS. This paper explores the social profile of unpaid carers and how this has evolved over time and place. It examines the demographic, socioeconomic and health characteristics influencing the likelihood of becoming an unpaid carer and the propensity to continue this role over time. It explores geographical differences in the propensity of becoming and/or remaining a carer and provides a longitudinal and place-based understanding of unpaid caregiving in England and Wales, informing policy and practice debates. We use data contained in the ONS Longitudinal Study from 2001 to 2021. Using binary logistic multilevel modelling to account for clustering within timepoints and geographical areas, it predicts probability of providing unpaid care, controlling for covariates (gender, age, ethnicity, household composition, marital status, care intensity, home tenure, education, economic activity, presence/number of people with long term illness, number of unpaid carers in household) and changes in time-varying covariates over time. Subnational analysis uses GOR and, for more localised analysis, the 1991 LADs, using harmonised geographies provided for ONS-LS to account for boundary changes over time.

Email: m.petrillo@sheffield.ac.uk

**A longitudinal analysis of Welsh and other language dynamics in England and Wales - preliminary results**  
**Oliver Duke-Williams - UCL, Bernice Kuang - University of Southampton, Scott Orford - University of Cardiff**

The ONS LS contains two questions about language use: one about ability to read, write speak or use Welsh, and another about main language used. The former is asked only in Wales, and is a long-standing question; the 2021 Census extends the series of responses, allowing further exploration of changes in Welsh language usage over time. The second was asked of all people, and was first included in the 2011 Census. The 2021 Census thus provides the first opportunity to look at longitudinal changes or stability in language used.

In this paper we present preliminary results from the 2021 census looking at both of these aspects of language. Whilst cross-sectional data shows a decline in Welsh speakers since 2011, we explore this as change over time, relating this to age, gender, migration and mobility within the UK. We also look at transitions in main language used, as related to individual demographic characteristics, as well as to self-declared proficiency in English.

Email: o.duke-williams@ucl.ac.uk

**National identity in transition?**

**Momoko Nishikido - Queens University Belfast, Stephen Jivraj - UCL, Gemma Catney - Queens University Belfast**

The 2011 Census shone a light on the divergence in feelings of national identity between ethnic majority and ethnic minority groups. White Britons living in England overwhelmingly described themselves as English, whereas ethnic minorities living in England overwhelmingly described themselves as British. White Britons living in Wales tended to describe themselves as Welsh rather than British, whereas ethnic minorities see themselves more as British rather than Welsh. The journey towards these identities will encompass individuals' ancestry, birthplace, language, culture, citizenship and feelings of inequality and discrimination. This project will provide evidence to support and develop theories of acculturation demonstrating how feelings of national identity varied by ethnic group in 2011 and 2021 and the stability between the two censuses. The central research questions will ask what drives ethnic minorities towards feelings of different forms of Britishness and how does living in an ethnically diverse neighbourhood affect national identity. The project is part of the ONS-LS Beta Test project to confirm data quality in the 2021 linkage.

Email: m.nishikido@qub.ac.uk

# Longitudinal census data – research using the newly linked 2021/2022 Census data 2

---

**Wednesday 3 September, 1pm**

## **The Longitudinal Impossible Dataset: Helping Users Navigate the ONS Longitudinal Study**

**Andreas Mastrosavvas, Nicola Shelton - University College London**

The Office for National Statistics Longitudinal Study (ONS-LS) follows a 1% sample of the population of England and Wales through each decennial Census, linking Census data with data from birth, death, and cancer registers. As one of the largest datasets of its kind in the UK, the ONS-LS is used for public good research on topics ranging from public health to labour market outcomes. However, access to the data is highly controlled and only possible via secure settings, meaning that researchers must often identify required variables and develop code prior to seeing the data. With thousands of variables available, navigating and exploring the available metadata can be a complex task for users.

This presentation will showcase the the Longitudinal Impossible Dataset (LIDS): a new customisable artificial dataset intended to familiarise prospective users with the data structures and variable domains represented in the ONS-LS.

Email: a.mastrosavvas@ucl.ac.uk

## **New 2021 Census questions - what have we learnt**

**Nicola Shelton - UCL, CeLSIUS**

The linked 2021 Census data from the ONS Longitudinal Study (the LS), is a 1% sample of individuals in the England and Wales Censuses. They share one of four birthdays. The data has been linked since 1971 and anyone born or arriving in England and Wales who has these birthdates is added to the data set. The data set is linked to cause specific mortality and cancer registrations. In 2021 new voluntary questions have been added for adults aged 16+ about sexual orientation and gender identity. There is also a new question asking about service in the UK Armed Forces.

This paper will share preliminary retrospective findings from beta test work on these questions.

## **Social Mobility and Health Outcomes of UK Armed Forces Veterans: Insights from the ONS Longitudinal Study**

**Umair Ali - Edinburgh Napier University, Carles Ibanez - Edinburgh Napier University, Alexandria Smith - Kings College, London, Iain Atherton - Edinburgh Napier University**

This study investigates how service in the UK Armed Forces shapes social mobility and health, while beta-testing the new census question pertaining to veteran status and its imputed values. We draw on the Office for National Statistics Longitudinal Study (ONS LS), which links a 1 percent representative sample of England and Wales across census waves from 1971 to 2021. First, we describe veterans' demographic and socioeconomic characteristics at the 2021 enumeration, including occupation, income proxies, and different measures of self-assessed health such as disability, and long-term illness indicators.

To examine social mobility, we reconstruct childhood socioeconomic status (age  $\leq 16$ ) from early census records and compare it with adult status in 2021. A matched-sampling design aligns veterans to non-veterans on age, sex, and socioeconomic status or region. This approach allows us to estimate how military service relates to upward or downward mobility.

A key component focuses on validity, in which we assess the completeness and accuracy of the imputed values of the service leavers question by comparing socioeconomic and health profiles of cases with and without direct veteran responses. By exploiting the longitudinal structure and occupation- and income-specific indicators in the ONS LS, we aim to map veterans' long-term mobility trajectories and explore which health disparities are explained by differences in childhood SES versus service experience.

The findings from this work support a deeper understanding of the census service-leaver question and offer a

robust, longitudinal portrait of veterans' social and health outcomes, informing policies to enhance their well-being.

Email: [u.ali@napier.ac.uk](mailto:u.ali@napier.ac.uk)

**Do upwardly mobile people live longer? An exploration of socioeconomic status and mortality using the ONS Longitudinal Study.**

**Merilynn Pratt, Louisa Blackwell, Jim Newman - Office for National Statistics**

The ONS Longitudinal Study links mortality outcomes and Census data from 1971 to 2021. It has been used since 1981 to describe and monitor inequalities in life expectancy, including between women and men, between those in different socioeconomic groups, and over time. In this paper we describe research on socioeconomic differences in mortality that uses the newly linked 2021 Census data. Having extra information on social class in 2021 provides new opportunities to study how social mobility affects a person's life expectancy.

LS-based studies of life expectancy by socioeconomic group have traditionally assigned social class at LS members' entry into the study. For children or adults where there is insufficient information to derive a socioeconomic group, they have been attributed a proxy, either from a parent or spouse. In this study, we derive socioeconomic group based on members' own classification at a later, post-entry census. We consider the impact of this new method on health improvement over time, and on the scale of observed inequalities.

We also consider social mobility trajectories for women and men between 2011 and 2021. We compare these trajectories to previous periods. We also compare the mortality outcomes between those who are socially mobile and those who remain stationary. Our intention is to use this study as a comparator for further research into mortality by other socioeconomic classifications, such as educational attainment and income.