UCL Institute of Education

Harmonised fertility histories in four British longitudinal cohort studies

CLS working paper number 2025/6

By Aase Villadsen, Samantha Parsons, Alice Goisis

CENTRE FOR LONGITUDINAL STUDIES

UKRI Economic and Social Research Council

Contact the author

Aase Villadsen

UCL Centre for Longitudinal Studies

a.villadsen@ucl.ac.uk

This working paper was first published in October 2025 by the UCL Centre for Longitudinal Studies.

UCL Social Research Institute

University College London

20 Bedford Way

London WC1H 0AL

www.cls.ucl.ac.uk

The UCL Centre for Longitudinal Studies (CLS) is an Economic and Social Research Council (ESRC) Resource Centre based at the UCL Social Research Institute, University College London. It is home to a unique series of UK national cohort studies.
For more information, visit www.cls.ucl.ac.uk.

This document is available in alternative formats. Please contact the Centre for Longitudinal Studies.
Email: clsdata@ucl.ac.uk

Disclaimer

This working paper has not been subject to peer review.

CLS working papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Any opinions expressed here are those of the author(s) and not those of the UCL Centre for Longitudinal Studies (CLS), the UCL Social Research Institute, University College London, or the Economic and Social Research Council.

How to cite this paper

Villadsen, A., Parsons, S., Goisis, A. (2025) *Harmonised fertility histories in four British longitudinal cohort studies* CLS Working Paper 2025/6. London: UCL Centre for Longitudinal Studies.

Abstract

Background

After World War II, Britain has witnessed significant societal changes, including in relation to fertility. Robust longitudinal and cross-cohort research requires data harmonisation to create comparable fertility measures to understand the predictors and consequences of these changes across generations.

Objective

This paper describes the variables included in newly created datasets containing harmonised fertility variables across four British longitudinal cohorts born between 1946 and 1990. The consistency with national statistics on fertility are examined.

Methods

The birth cohorts are: National Survey of Health and Development born in 1946, the 1958 National Child Development Study, the 1970 British Cohort Study, and Next Steps born in 1989/90.

Results

The harmonised datasets include information on the cohort members' biological children at each survey sweep during childbearing age, such as whether they have any and, if so, how many, age of eldest and youngest child, number of boys and girls, and children by a previous partner. Additional variables related to non-biological children have been derived where possible. The percentage of female cohort members who have had at least one live birth and the number of children at each five-year interval from age 20 to 50 is highly consistent with national statistics on fertility.

Conclusion

The harmonised fertility datasets on four birth cohorts in Britain improve the measurement of fertility in cohort analyses and these data show high consistency with national statistics on fertility.

Keywords: Fertility, UK cohort studies, harmonised data, cross-cohort analyses

1. Introduction

Literature /background

Since World War II, Britain has experienced significant changes in fertility (Berrington et al., 2015, Ermisch, 2021), marked by the post-war baby boom, declining birth rates and the postponement of childbearing from the 1970s. Advances in contraception, shifting gender roles, increased education, changes in social norms, economic uncertainty, and evolving family structures have contributed to these trends (Mills et al., 2011). Yet, many questions remain about the causes and consequences of these important demographic shifts and of low fertility levels (Graham, 2021). High quality data on fertility harmonised across several datasets in Britain can contribute to understand the predictors and consequences of these changes over time and across generations.

Aims of paper

Britain disposes of a unique series of national birth cohort studies which follow the lives of multiple generations of people and collect rich data on multiple domains across the life course, including fertility. Through interviews at different ages, cohort members report on each child they have had, including the date of birth, sex, and year of birth. However, in their raw format these data at the child level are not straightforward for researchers to use when, for example, they are looking for a summary measure of the cohort members' fertility, such as whether they have had any children, at what age and how many across their entire reproductive lifespan.

The primary aim of this paper is to describe newly created datasets on fertility histories that have been derived and harmonised across four British cohort studies born in 1946, 1958, 1970 and 1989/90, and which is now available to researchers (UCL Centre for Longitudinal Studies., 2025). The secondary aim of this paper is to examine the consistency between the birth cohort data and official statistics on fertility published by the Office for National Statistics.

2. Data and Methods

The cohort studies

The data were collected as part of four separate cohort studies born in Britain between 1946 and 1989/90, which have followed large samples of individuals across time, most since their birth. The cohorts are: MRC National Survey of Health and Development (1946), 1958 National Child Development Study, 1970 British Cohort Study, and Next Steps (1989/90). Most interviews across these cohorts were conducted face to face, but with some sweeps involving postal or telephone surveys, and some using mixed modes (see data guide for details, Villadsen et al. (2025)). Amongst a wide and extremely rich collection of measures around the cohort members lives (spanning from family socio-demographic characteristics, cohort members education, physical and mental health and cognition), the studies collect information on cohort members' fertility. These large and well-known cohorts have contributed significantly to research across the scientific community and to key policy areas. The following will provide a brief description of each of these cohort studies.

MRC National Survey of Health and Development study (NSHD). This is the oldest and longest running of the British birth cohort studies. From an initial maternity survey of 13,687 of all births recorded in England, Scotland and Wales during one week in 1946, a socially stratified sample of 5,362 singleton babies born to married parents was selected for follow-up. During their childhood, the main aim of the NSHD was to investigate how the environment at home and at school affected physical and mental development and educational attainment. During adulthood, the main aim was to investigate how childhood health and development and lifetime social circumstances affected their adult health and function and how these change with age. The sweeps included in the NSHD fertility dataset span eight sweeps from age 19 to age 53 (collected between 1965 to 1999). The sample sizes across these sweeps range from 3,034 (males=1,471, females=1,563) to 3,897 (males=2021, females=1,876).

1958 National Child Development Study (NCDS). The NCDS follows the lives of over 17,000 people born in England, Scotland, and Wales in a single week in March of

1958. The first survey was carried out at the birth of the child, with subsequent follow up surveys at age 7 through to age 62. The initial survey was known as the 'Perinatal Mortality Study' with a focus on factors associated with the health of mothers and their newborn babies. The survey sweeps included in the NCDS fertility dataset included five sweeps from age 23 to 50 (collected between 1981 and 2008). Sample sizes range between 9,532 (males=4,643, females=4,889) and 12,535 (males=6,266, females=6,269) cohort members.

1970 British Cohort Study (BCS70). The BCS follows the lives of around 17,000 people born in England, Scotland, and Wales in a single week in April of 1970. The initial survey was carried out at the birth of the child, with follow up surveys at age 5, 10, 16, 26, 30, 34,38, 42, 46, and 51 years. As a longitudinal and multidisciplinary study, a large amount of varied and rich information has been collected about cohort members through their lives. Seven sweeps from the BCS from age 26 to 51 were included in the fertility dataset (collected between 1996 and 2021). Sample sizes range from 8,016 (males=3,802, females=4,210) to 11,260 (males=5,471, females=5,789) cohort members.

Next Steps. Previously known as the Longitudinal Study of Young People in England (LSYPE), this is the youngest cohort study for which fertility variables are derived. It began following around 16,000 young people in state or independent schools across England who were in Year 9 in 2004, born mainly in 1989 and 1990. Since the initial survey when cohort members were around age 14, there were annual follow up surveys until age 20, and after this at age 25 and 32. The initial focus on the study was on cohort members' schooling and their further transition into education and employment. The data used in the Next Steps fertility dataset are from surveys at age 25 and age 32 (collected in 2015 and 2022), with samples sizes of 7,705 (males=3,427, females=4,278) and 7,279 (males=3,154, females=4,125), respectively. Because of the younger age of this cohort, we are unable to construct complete fertility histories. As more survey sweeps becomes available information these variables can be constructed.

Variables in datasets

The variables included in the harmonised fertility datasets are listed in Box 1. We used the approach of deriving as many of these as possible in each cohort and their respective sweeps, rather reducing the list to the lowest common denominator. The NSHD especially was affected by lack of information (particularly on non-biological children), whereas a larger number of harmonised variables could be derived in the NCDS, BCS, and Next Steps, which shared many identical survey questions. There were four main groupings of target variables, 1) those relating to biological children (live births), 2) non-biological children, 3) all children (biological and non-biological), and 4) children combined with partnership status. Additional variables were included, such as research case ID (to enable linkage to other variables from these cohorts), sex of cohort member, sweep participation, and weights to adjust for survey design in cohorts where this is applicable.

We created a separate longitudinal dataset for each cohort, with data covering survey sweeps from when cohort members were young adults generally in their early to mid-twenties, until early fifties which for most marks the end of the reproductive window. The derived variables provide a summary of fertility (live births) for each cohort member at each sweep, such as whether the cohort member had had any children, number of children, age of the eldest and youngest child, and number of boys and girls. The focus was on live births rather than pregnancies that were terminated, miscarriages or stillbirths. Further details on how the fertility variables were derived in each of the cohorts can be found in the data user guide (Villadsen et al., 2025) In addition, annotated Stata code for variable derivation is available for all cohorts on GitHub, which allows data users to inspect the derivation process, as well as adapt the code for creation of further and alternative fertility variables.

Examining fertility in the cohorts

For each cohort, we show the proportion of males and females who have given birth to or fathered a child at five-years intervals from age 20 to 50. We focused on live births and excluded still births. Although cohort members were not always interviewed at these specific ages, we are able to work out at each age whether they

have had a child by using the derived variable on the age of their eldest biological child. By using the same age points, we can compare fertility trends across generations. Sampling weights (see sampling details above) are applied to all analyses of NSHD as available across all survey sweeps. Weights are not available in the NCDS; for BCS, attrition weights are only available at age 51 and are therefore not applied to the analyse to retain consistency between sweeps. In Next Steps, a combined sampling and attrition weight is available at both sweeps and used on our analyses.

Comparison to national childbearing statistics

To quality check the fertility variables in our datasets, we present comparisons between our estimates and the national statistics on childbearing (any children and number of children) for all females in England and Wales born in the same years as our cohort members (Office for National Statistics, 2024). Similar statistics are not available for males (Dudel and Klüsener, 2019) so these comparisons are limited to female cohort members.

3. Results

Proportion of cohort members who have children

Table 1 shows the percentage of females and males in each cohort who have given birth to/fathered at least one live birth by five years age intervals from age 20 to 50.

We see a clear trend that fertility is postponed towards older ages in the younger cohorts compared to the older. By age 30, 76.5% (95% CI: 74.0-78.9) of NSHD females had started a family, dropping in subsequent cohorts to 70.2% (69.0-71.4) in NCDS, 58.2% (56.8-59.5) in BCS, and 54.7% (52.8-56.5) in Next Steps. For males we see a similar postponement pattern across cohorts, but within each cohort we also see that males tend to start their families later than females. At age 30, 65.3% (62.5-68.1) of NSHD males had at least one child, dropping to 57.4% (56.0-58.7) in the NCDS, 42.7% (41.2-44.1) in BCS, and 40.1% (37.9-42.3) in Next Steps.

At the end of the reproductive period at age 50, a higher proportion of NSHD females had a least one child 87.2% (85.2-89.1), compared to 83.5% (82.5-84.6) of NCDS females, and for BCS females this was 84.2% (83.1-85.3). In other words, compared to the females born in 1946 in the NSHD, a decreasing number of females in the successive cohorts in 1958 (NCDS) and in 1970 (BCS) has at least one child. The same pattern was largely observed for males; at age 50, 82.7% (80.5-85.0) of NSHD males had fathered a child, falling to 79.6% (78.4-80.7) in the NCDS, and 81.3% (80.0-78.5) in the BCS. Compared to females, a statistically significantly larger proportion of males had no children by age 50, which was a consistent pattern within all cohorts.

Number of children of cohort members

In Figure 2 we show the number of live biological children cohort members had given birth to or fathered by each survey sweep in each respective cohort. Figures are presented for males and females, along with national statistics for England and Wales for females of the same age born in the same year as the cohort members. At the end of fertility around age 50, we see that the number of children decreases as

cohorts get younger. In the NSHD, the average number of children born to females was 2.01 (95% CI: 1.95-2.08), and to males 1.90 (1.82-1.97), falling to 1.92 (1.89-1.96) for females and 1.81 (1.77-1.84) for males in the NCDS, and then decreasing in the BCS for females to 1.86 (1.82-1.90) whilst increasing marginally for males to 1.85 (1.81-1.90). Across all cohorts we see that females tend to have slightly more children than males, although only the sex difference in the NCDS was fully statistically significant.

How our data compare to national statistics on fertility rates

Figure 1 shows the proportion of women with at least one live birth in the cohorts and in national statistics by age. We generally see a good degree of consistency between the two, although with a tendency for the cohort figures to track slightly above those of national figures at most ages. Figure 2 shows the number of children born to women in the cohorts and in national statistics, by age. Again, figures are largely consistent, although there is a tendency for female cohort members to have slightly fewer children than in the national data, especially in the BCS, but also in the NSHD and NCDS in the later sweeps.

Methodological differences between national statistics and the cohorts might contribute to the small differences we observe. National statistics include all women in England and Wales regardless of their birthplace in the UK or abroad, whilst the cohorts are all born in Wales, England, or Scotland - with exception of Next Steps who were recruited around the age of 14 and 15 in England but could be born elsewhere in the UK or abroad. Attrition in the cohorts may also play a role, affecting especially respondents with lower socioeconomic status (Silverwood, 2024) who tend to have a higher level of fertility (Berrington et al., 2015). Indeed, in additional analyses in which we applied the available attrition weights in the BCS at age 51 (the only sweep which includes weights in BCS70), we see a higher number of biological children for females (2.02) than the unweighted number (1.86). Similarly, in Next Steps where a combined sampling and attrition weight is available, analyses that omit the attrition weight (making it comparable to the other cohorts) result in lower number of children born to females both at age 25 (0.43) and at age 32 (1.05),

compared to attrition weighted results at age 25 (0.57) and age 32 (1.19). See footnote in Figure 2.

Data availability and access

Fertility data for the cohorts can be accessed via the UK Data Service (UKDS). Users need to register with the UKDS, and data for NCDS, BCS, and Next Steps can be downloaded once the End User Licence has been accepted by the user. The NSHD fertility data can be downloaded once the Special Licence application form has been reviewed by UKDS and approved by the NSHD Data Sharing Committee. For the NCDS, BCS, and Next Steps, other data from these cohorts are also available via the UKDS and can be easily linked to the harmonised fertility datasets using individual IDs. Those wishing to obtain additional variables for the NSHD alongside fertility data, need to apply for access via the MRC Unit for Lifelong Health and Ageing at UCL (LHA), which manages the NSHD. We refer to the data user guide for these datasets for further information (Villadsen et al., 2025).

Conclusion

The harmonised fertility datasets on four birth cohorts in Britain improve the measurement of fertility (as an outcome, predictor, or control variable) in within cohort analyses, and are well placed to facilitate cross-cohort research on fertility such as understanding the drivers and consequences of changes over time in the timing and number of children born to females and males. Reassuringly, these data show high consistency with national statistics on fertility.

Acknowledgements

This work was undertaken as part of the Centre for Longitudinal Studies Resource Centre funded by the ESRC (ES/W013142/1). We would also like to thank our colleagues at Unit for Lifelong Health and Ageing at UCL for their collaboration around the NSHD data.

Box 1: List of target fertility variables for each sweep in each of the cohorts

CM biological children

Whether has had any bio children

Number of bio children

Flag: More biological children reported at previous age than at current age

Number of bio children in HH

Flag: More bio children reported in HH grid than in pregnancy data

Number of bio children not in HH

Number of bio children had with a previous partner

Have had any bio children with a previous partner

Age in years of eldest bio child

Age in years of youngest bio child

Age in years of CM at birth of eldest bio child

Age in years of CM at birth of youngest bio child

Number of bio children who are boys

Number of bio children who are girls

CM non-biological children

Whether has any non-bio children in HH

Number of non-bio children in HH

Number of adopted children in HH

Number of fostered children in HH

Number of stepchildren in HH

Age in years of eldest non-bio child

Age in years of youngest non-bio child

Number of non-bio children who are boys

Number of non-bio children who are girls

CM biological or non-biological children

Whether has any children (bio or non-bio)

Number of children (bio or non-bio)

Age in years of eldest child (bio or non-bio)

Age in years of youngest child (bio or non-bio)

Number of children who are boys (bio or non-bio)

Number of children who are girls (bio or non-bio)

CM partner and/or children

Whether has a partner in HH

Marital status

Whether has live-in partner/spouse partner and/or any bio children

Whether has live-in partner/spouse partner and/or any bio or non-bio children

Other variables

CM research case identifier

Sex of CM

Birth year of CM

Birth month of CM

Whether CM took part in survey sweep

Interview year

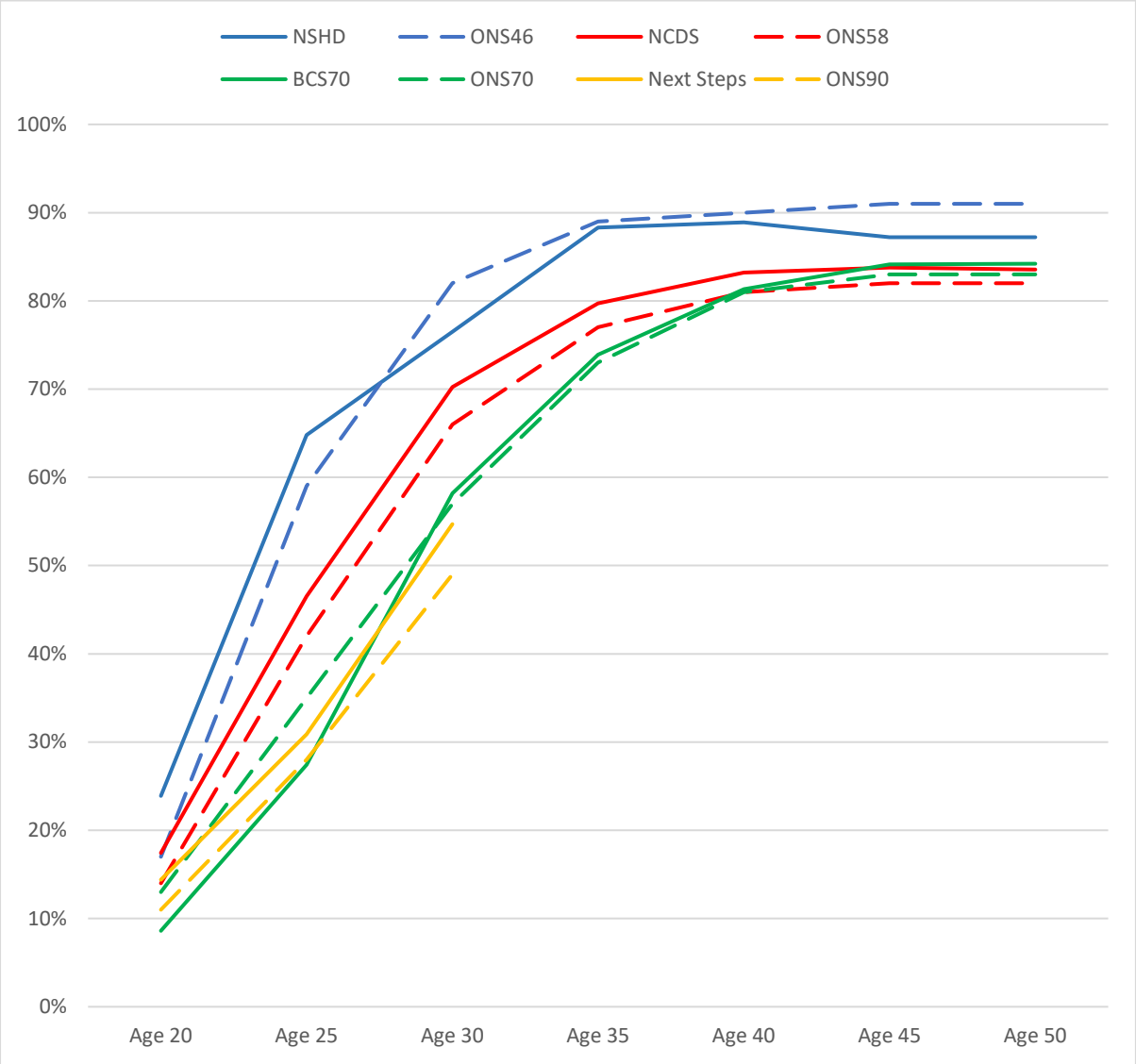
Interview month

Survey weights (in applicable cohorts)

Table 1: Percentage of females and males with at least one live birth by age

	Sample size	Males			Females		
		Prop	95% CI low	95% CI high	Prop	95% CI low	95% CI high
NSHD							
Age 20	3,749	8.3%	6.7%	9.8%	23.9%	21.5%	26.3%
Age 25	3,749	43.4%	40.7%	46.2%	64.8%	62.2%	67.4%
Age 30	3,339	65.3%	62.5%	68.1%	76.5%	74.0%	78.9%
Age 35	3,321	79.9%	77.7%	82.2%	88.3%	86.5%	90.1%
Age 40	3,261	81.9%	79.8%	84.1%	88.9%	87.1%	90.7%
Age 45	3,034	82.3%	80.0%	84.6%	87.2%	85.2%	89.1%
Age 50	3,034	82.7%	80.5%	85.0%	87.2%	85.2%	89.1%
NCDS							
Age 20	12,528	6.1%	5.6%	6.8%	17.4%	16.5%	18.4%
Age 25	11,186	30.0%	28.8%	31.2%	46.5%	45.2%	47.8%
Age 30	11,186	57.4%	56.0%	58.7%	70.2%	69.0%	71.4%
Age 35	11,173	70.2%	69.0%	71.4%	79.7%	78.7%	80.8%
Age 40	11,173	76.2%	75.1%	77.3%	83.2%	82.2%	84.2%
Age 45	9,343	78.9%	77.7%	80.1%	83.8%	82.7%	84.8%
Age 50	9,430	79.6%	78.4%	80.7%	83.5%	82.5%	84.6%
BCS							
Age 20	8,210	2.3%	1.9%	2.8%	8.6%	7.8%	9.5%
Age 25	8,210	13.4%	12.4%	14.6%	27.4%	26.2%	28.7%
Age 30	9,578	42.7%	41.2%	44.1%	58.2%	56.8%	59.5%
Age 35	8,776	63.5%	62.0%	65.0%	73.9%	72.6%	75.2%
Age 40	9,664	74.8%	73.5%	76.0%	81.3%	80.2%	82.4%
Age 45	8,145	78.0%	76.7%	79.3%	84.1%	83.0%	85.2%
Age 50	7,674	81.3%	80.0%	82.5%	84.2%	83.1%	85.3%
Next Steps							
Age 20	7,624	5.0%	4.1%	6.2%	14.4%	13.0%	15.9%
Age 25	7,624	16.9%	15.3%	18.7%	30.9%	29.1%	32.7%
Age 30	6,683	40.1%	37.9%	42.3%	54.7%	52.8%	56.5%

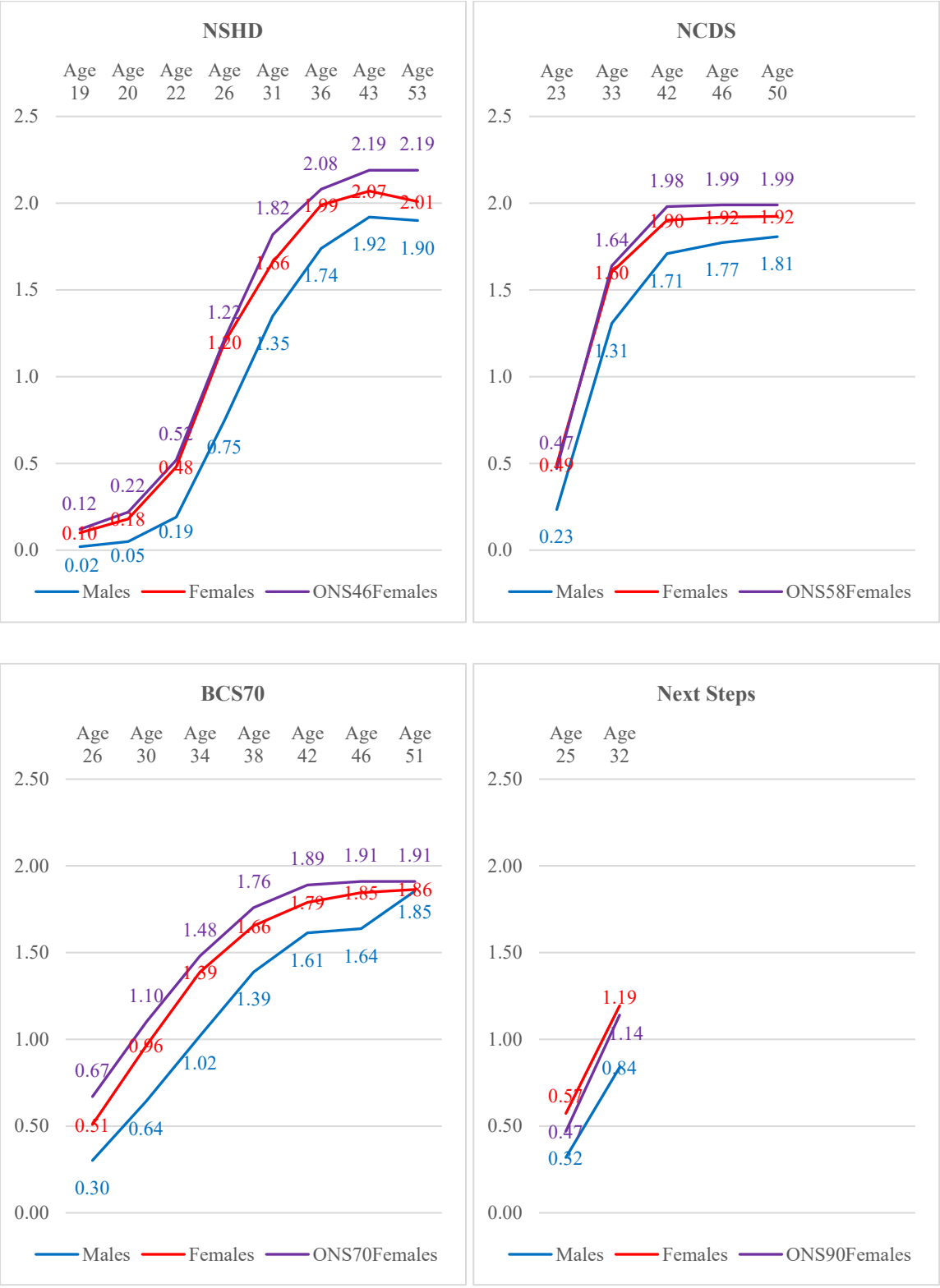
Figure 1: Percentage of females with at least one live birth by age – comparing cohort data to ONS national statistics



Note: Within each cohort sample sizes vary across age points due to attrition over time.

ONS statistics are published figures from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/conceptionandfertilityrates/datasets/childbearingforwomenbornindifferentyearsreferencetable

Figure 2: Number of biological children, by age and birth cohorts



Note: Within each cohort sample sizes vary across age points due to attrition over time.

ONS statistics are published figures from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/conceptionandfertilityrates/datasets/childbearingforwomenbornindifferentyearsreferencetable

In the BCS an attrition weighting variable is available at age 51, although this has not been used in analyses to be consistent with the previous sweeps for which no weight is available. Weighted results for BCS at age 51 showed a slightly higher number of children born to females (2.02) and fathered by males (1.94) than the unweighted number of children shown in the figure.

In Next Steps weight a combined sampling and attrition weight is available at both age 25 and 32 and therefore applied to analyses in the figure. However, without the attrition weight but retaining the survey design weight, we see a lower number of children born to cohort members (age 25: males 0.24, females 0.43; age 32: males 0.76, females 1.05).

References

Berrington, A., Stone, J. & Beaujouan, E. 2015. Educational differences in timing and quantum of childbearing in Britain: A study of cohorts born 1940–1969. *Demographic Research,* 33**,** 733-764.

Dudel, C. & Klüsener, S. 2019. Estimating men's fertility from vital registration data with missing values. *Population Studies,* 73**,** 439-449.

Ermisch, J. 2021. English fertility heads south. *Demographic Research,* 45**,** 903-916.

Graham, E. 2021. Theory and explanation in demography: The case of low fertility in Europe. *Popul Stud (Camb),* 75**,** 133-155.

Mills, M., Rindfuss, R. R., Mcdonald, P. & Te Velde, E. 2011. Why do people postpone parenthood? Reasons and social policy incentives. *Human reproduction update,* 17**,** 848-860.

Office for National Statistics. 2024. *Childbearing for women born in different years* [Online]. Newport, Wales: Office for National Statistics. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/conceptionandfertilityrates/datasets/childbearingforwomenbornindifferentyearsreferencetable [Accessed 21 Jan 2025].

Silverwood, R. 2024. Handling missing data in the CLS cohort studies. London: UCL Centre for Longitudinal Studies.

UCL Centre for Longitudinal Studies. 2025. Harmonised Fertility Histories in the National Child Development Study, 1970 British Cohort Study and Next Steps, 1981-2024. [data collection]. UK Data Service.

Villadsen, A., Parsons, S. & Goisis, A. 2025. Fertility histories in four British cohort studies. London: UCL Centre for Longitudinal Studies.