UNIVERSITY OF SOUTHAMPTON

FACULTY OF SOCIAL AND HUMAN SCIENCES

Social Sciences

Estimating the fertility of migrants to England and Wales using the Office for National Statistics Longitudinal Study

by

James William Robards

Thesis for the degree of Doctor of Philosophy

March 2012

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL AND HUMAN SCIENCES

Social Sciences

Doctor of Philosophy

ESTIMATING THE FERTILITY OF MIGRANTS TO ENGLAND AND WALES USING THE OFFICE FOR NATIONAL STATISTICS LONGITUDINAL STUDY

by James William Robards

Since 2001, there has been a consistent year-on-year increase in the period total fertility rate for England and Wales. At the same time migration to England and Wales has accelerated from the late 1990s. It is possible that the large number of migrants of childbearing ages moving to England and Wales, larger family size norms among foreign born women and a birth timing effect among recent migrants to England and Wales have led to the increase in the total fertility rate. However, the relative influence of any timing effect among recent migrants on the total fertility rate is not known. Research on migrant fertility in France (Toulemon, 2004) and Sweden (Andersson, 2004) has identified elevated fertility among migrants in the time period immediately after the migration event. Conversely, research in England and Wales has focused on period measures of fertility rather than estimating whether there is an elevated level of fertility among the large number of recent migrants to England and Wales. The first aim of this thesis is to accurately account for non-continually resident members of the Office for National Statistics (ONS) Longitudinal Study (LS) between census dates and use these LS members in fertility analysis. The second aim of this thesis is to investigate whether migrants to England and Wales show an elevated level of fertility after migration. It is only possible to estimate the fertility of recent migrants provided the sample exposed to risk of giving birth can be identified.

Contents

List of Tables

List of Figures

List of Appendices

Declaration of Authorship

I, James William Robards, declare that the thesis entitled *Estimating the fertility of migrants to England and Wales using the Office for National Statistics Longitudinal Study* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

Signed: . . . .

Date March 2012

Acknowledgements

This work was completed at the University of Southampton with the financial support of the Economic and Social Research Council (ESRC), studentship number ES/G018766/1 without which this would not have been possible.

I am very grateful for the continued support, guidance and encouragement of my supervisors Andrew Hinde and Ann Berrington.

I wish to thank Jakub Bijak (University of Southampton) and Michael Rendall (Maryland Population Research Center) who examined this thesis at the University of Southampton in May 2012.

This research has been supported by Chris Marshall, Rachel Stuchbury and the staff at the Centre for Longitudinal Study Information and User Support (CeLSIUS) who provided user support in the form of an extract of and access to the Office for National Statistics (ONS) Longitudinal Study (LS).

The Longitudinal Study Development Team at the ONS Titchfield office assisted with access to the dataset in a secure Virtual Microdata Laboratory (VML) setting. Their assistance with this access to the data and the discussions on the workings of the LS are much appreciated as is feedback from Steve Smallwood and Julie Jeffries on presentations made at the ONS. The Micro-data Analysis and User Support branch at ONS Newport cleared intermediate and final outputs from this research.

My employers at the University of Southampton have supported me in the latter part of the preparation of this thesis. I wish to thank Maria Evandrou, Jane Falkingham, Athina Vlachantoni and colleagues in the Care Life Cycle and Centre for Population Change research clusters.

I wish to thank my parents, Tony and Hazel and my sister, Sarah for their support and encouragement.

I also wish to acknowledge the on-going support of family and friends; Sylvia Groundsell, Peter Groundsell, Georgia Chronaki and Matt Loxham, William Dawes, Anton Blackie, Claire Blackie, Mark Lyons-Amos, Amos Channon, Melanie Frost, Robert McIlveen, Malcolm Campbell, Paul Hayward, Peter Barling, Tom Ritman-Meer, William Carr, Sian Brett, Hollie Curtice, David Culliford, Mike Blewitt, Guy Abel, Claire Cheng, Sophie Dear, Tammy Miller, Heleen Vanneste, Eleni Tzortzi, Matthias Feinäugle, Victor Bondanese, Eirini Vavatsikou, Karolina Vavatsikou, Mark Knowles, Kim Cartwright, Tim Bridger and Samantha Shave.

The research environment at the School of Social Sciences has proved inspirational and been perfect for the completion of this work. I wish to thank the support of my fellow post-graduates past and present including; Maryam Alriyamy, Nayef Alsadoun, Jennifer Baird, Bernard Bauffour, Reza Boostani, George Disney, Marcus Green, Corrado Giulietti, Clare Woodford, Derya Tash, Greg Taylor, Alessandro Mennuni, Olga Maslovskaya, Kate Jowett, Thomas King, Alex Skew, Jana Javornik-Skrbinsek, Elif Kara, Luis Pinedo, Claire Bailey and anyone else who in some form helped with the successful realisation of thesis.

Peter Jackson, Charles Pattie, Deborah Sporton and Dan Vickers at the University of Sheffield and Chris Simpson, Vicki Popplewell, Robert Hack and Paul Jeffrey at ECORYS supported me in my earlier studies and encouraged me to pursue this PhD at the University of Southampton.

Publications from this thesis

Material from Chapters 4 and 5 of this thesis were published in the June 2011 issue of *Population Trends* from the Office for National Statistics.

Robards, J.W., Berrington, A. and Hinde, A. (2011) Estimating fertility rates using the ONS Longitudinal Study – what difference does the inclusion of non-continually resident members make? *Population Trends*, 144, 33-47.

Office for National Statistics Longitudinal Study – cell counts and Crown copyright

The permission of the Office for National Statistics to use the Longitudinal Study is gratefully acknowledged, as is the help provided by staff of the Centre for Longitudinal Study Information & User Support (CeLSIUS). CeLSIUS is supported by the ESRC Census of Population Programme (Award Ref: RES-348-25-0004). The author alone is responsible for the interpretation of the data. Clearance number 30106.

A note on cell counts / anonymity of LS members

In some tables small cell counts have led to the removal of values from tables. The Office for National Statistics does not allow the publication of cell counts of less than '10'. Where values have been removed this is shown by a * and is a footnote has been included.

A note on National Health Service terminology used in this thesis

Since this thesis was started changes in responsibility for the National Health Service Central Register have led to adjustments in some terminology. In the thesis the superseded versions of the terminology, consistent with Office for National Statistics Longitudinal Study publications have been used.

The National Health Service Central Register (NHSCR) is now called the National Health Service Information Centre (NHSIC). This came into effect on 1 April 2008 when responsibility for the NHSCR moved from the Office for National Statistics to the NHS.

The National Health Service (NHS) Central Health Register Inquiry System (CHRIS) is now called the Patient Demographic System (PDS).

Glossary

CeLSIUS - Centre for Longitudinal Study Information and User Support. CeLSIUS makes the Office for National Statistics Longitudinal Study (LS) available for academic research.

Embarkation - refers to recorded 'exits' from the LS. These exit events come in the form of migrations from England and Wales, de-registration with an National Health Service (NHS) General Practitioner (GP) or 'cancellations' for those who are registered with a GP and when asked, fail to confirm that they are still resident.

Filter - the procedure in SPSS for selecting a sample for analysis. Members in the sample selected meet the criteria that is detailed in syntax.

Matching - the process of finding an existing LS record for a person born on an LS birth date and resident at a census. The measurement of the longitudinal matching of LS members between census dates is done through the calculation of linkage rates (see below).

Linkage rate - calculated by the proportion of members of a given group (e.g. age in census year) who are successfully matched at a census to an LS record at the last census. 'Valid exits' in the form of deaths or recorded embarkations (via de-registration at an NHS GP) are taken into account in the calculation of linkage rates. Individuals are not identified (matched) if they were not at a census, information on their death or embarkation was not recorded on the National Health Service Central

Register (NHSCR) or there are inconsistencies between dates of birth, or other personal information given at censuses and used for matching.

NHSCR - National Health Service Central Register. The NHSCR is part of the Office for National Statistics and compiles a record for each NHS patient who is registered with an NHS General Practitioner in England, Wales or the Isle of Man. The actual database used in extracting data for the LS is the Central Health Register Inquiry System (CHRIS). A record for almost every member of the population is held on the NHSCR and events (e.g. cancer registrations) are routinely recorded on the database which made it attractive for use by the LS.

Sampling fraction - the degree to which the LS represents the actual population as calculated from vital statistics. Sampling fractions are calculated using traced LS sample members and census populations without adjustment for census under-enumeration (see below). The LS sample is divided by vital statistics data or mid-year population estimates.

Tracing - identifying an LS member on CHRIS in Southport by automatic or manual methods. This is done at the NHSCR using data from the census provided by the LS team at Titchfield. The NHSCR enables records for LS members to be linked to various life events for these individuals. It also facilitates matching of records collected at different points.

Tracing rate - Tracing rates indicate the likelihood of both census and event data linkage for groups within the Longitudinal Study. Potential LS members (persons with an LS birth date) are 'untraced' on the NHSCR if they have not been found because they are not registered with a doctor, or inconsistent names or dates of birth have been used.

Under-enumeration - refers to an undercount of persons or households in the census. Census under-enumeration must be understood in the context of the 'One Number

Census' which was adopted in 2001. The full census was completed, after which an independent follow-up Census Coverage Survey was conducted. This coverage survey was stratified by a 'hard to count' index based upon characteristics likely to be associated with under-enumeration, such as the number of multi-occupied addresses. Using both the raw data and the coverage survey the total resident population was estimated (one number) and synthetic people could be imputed into synthetic households or enumerated households so that the Census database was fully adjusted for biased under-enumeration at a local level. Because the LS is concerned with the accurate tracing of individuals through time and the use of NHS records for the attachment of information the LS uses only 'real' persons from the census and not any imputed cases. This means that there are no imputed LS members but missing responses to questions at the 2001 census led to the imputation of values. These are imputed responses to item non-response in census variables. Where imputed variables have been used they are identifiable.

Valid exit - an embarkation or death which has been recorded on the NHSCR and therefore has a date and order number if there has been more than one occurrence of this event.

Chapter 1

Introduction

This PhD thesis consists of nine chapters of research on the fertility of recent migrants using the Office for National Statistics Longitudinal Study for England and Wales. Since 2001 there has been a persistent year-on-year increase in the total fertility rate, at the same time as rising migration to England and Wales. Increases in migration occurred because of geopolitical instability (i.e. conflict in the Balkans, the Democratic Republic of the Congo, Sierra Leone, Afghanistan and Iraq) and on-going processes of globalisation in education and employment. After the accession of eight Eastern Central European countries (along with Malta and Cyprus) to the European Union in 2004 migration levels accelerated further.

Migration and fertility are interrelated events, with migration impacting on fertility in at least three ways. Firstly, migrants tend to be within the key childbearing age ranges and therefore expand the population of women exposed to risk of birth. Between 2001 and 2007 the proportion of total births to women born outside of the UK rose from 15% to 22% (Tromans et al., 2009). A geographical match between local government areas where there has been a substantial increase in fertility, and areas where there has been a high rate of migration between 1986 and 2006, has also been identified (Tromans et al. 2008). Secondly, some migrant groups tend to have higher completed family sizes than the non-migrant population, Pakistani and

Bangladeshi women continue to have higher completed family sizes than UK born women (Dunnell, 2007) and are more likely to migrate for family formation (Peach, 2006). Thirdly, period fertility rates can be inflated by a timing effect among recent migrants propensity to give birth shortly after migration (Andersson, 2004; Toulemon, 2004). In England and Wales the first two of these drivers of change within the migrant-fertility relationship have been the subject of investigation, but there has been no work looking at the timing of fertility among recent migrants. Therefore, the key aim of this thesis is to investigate whether migrants to England and Wales show an elevated level of fertility shortly after their migration. The second main aim of this thesis is to accurately account for non-continually resident LS members between census dates and use these LS members in fertility analysis. Frequently, fertility analysis using the LS has used a sample of continually resident LS members. This research wishes to use the LS in a more dynamic manner to understand what the implications are for the measurement of fertility with the inclusion of non-continually resident LS members. The chapters in this thesis build incrementally towards the main aim, paying careful attention to sample selection and the sample of women from the ONS LS accurately exposed to risk of birth.

As already referred to, there are two key works by Andersson (2004) and Toulemon (2004) who, among others (Ford, 1990; Goldstein and Goldstein, 1981; Mussino and Strozza, 2012; Mussino et al., 2009; Schrool, 1990), have identified an elevated level of fertility immediately after the migration event. Toulemon (2004) identifies that very high fertility among migrants to France immediately after arrival declines to match ‘French’ fertility quite quickly. The approach which Toulemon takes allows the calculation of fertility rates by the length of time since arrival and considers the fertility history of the migrant. Meanwhile, Andersson (2004) used Swedish registry data to study the migration-fertility relationship. In their first two years of residence in Sweden, first birth rates for migrants are 100% higher than those of Swedish women. It is concluded that migration and family-building are inter-related processes. In Chapter 2 the relationship between migration and subsequent fertility is fully explained with a range of material researching this discussed. Given the high migration observed, it is tenable that a timing preference among the large number of migrants to England and Wales has impacted on the TFR.

Data for the collection of a suitably sized sample of migrants and one which records the desired socioeconomic information for event history analysis is scarce (Sigle-Rushton, 2008). In England and Wales the Office for National Statistics (ONS) Longitudinal Study (LS) represents an approximate 1% sample of the population of England and Wales using a combination of census data, National Health Service Central Register (NHSCR) data and Vital Registration data (births and deaths) (Hattersley and Creeser, 1995). New LS members enter the data through the NHSCR and because of the way in which this dataset is a study composed of routinely collected data, there is not the response burden of repeated surveying typical of some other longitudinal datasets (Blackwell et al. 2005). Although the LS has been used in previous fertility research in England and Wales (Ekert-Jaffé et al., 2002 ; Portanti and Whitworth, 2009; Rendall, 2003; Rendall et al., 2005; Rendall et al., 2009; Rendall and Smallwood, 2003; Werner, 1988), there has been a lack of research on how the functioning of the LS and the sample of which it is composed changes through time and hence makes the data more appropriate for studying some members of the population than others. Therefore, in Chapter 2, research which has used the LS is discussed before Chapter 3 fully explains the data of which the LS is composed, the functioning of the dataset and past research approaches to sample selection.

Fully appreciating the functioning of the ONS LS and past research approaches using the dataset is necessary because some past research using the LS has provided insufficient information on the characteristics of the samples selected and the ways in which the sample might not be representative of the national population. In particular, through the use of cohorts which were resident at successive census dates (e.g. completing a census form at 1981, 1991 and 2001), there is the potential to introduce bias by virtue of the cumulative socioeconomic differences in the linkage of individuals between census dates. Therefore, Chapter 4 is concerned with identifying a sample of female LS members for whom there is complete information on their whereabouts and residence throughout the 1991-2001 and 2001-2007 periods. These LS members are identified through the creation of two types of LS member residence trajectories – 'consistent cases' (where all information on residence is coherent and complete (as recorded in the dataset)) and 'inconsistent cases' (those cases where there is some form of missing data for the LS member or an illogical / non-matching set of events). Among the inconsistent cases the most numerous subset are those where there was attrition between 1991 and 2001. Therefore, later in the chapter the relative weight of different socioeconomic characteristics in attrition in the ONS LS between 1991 and 2001 are identified. For selection of an appropriate sample for event history analysis this is an important stage in the analysis.

Given that this thesis is concerned with the fertility and identifying the sample exposed to risk of birth, Chapter 5 looks at the fertility rates which can be derived from the ONS LS and how these compare to official statistics. There has been no recent work exploring the comparability of fertility rates from the ONS LS with those published by the ONS. With the residence trajectories identified in Chapter 4, the identification of the samples of women and corresponding births in non-census years is completed with a higher degree of accuracy and certainty than would otherwise be the case. Fertility rates for single years between 1991 and 2001 are calculated before the denominator and numerator are compared to the corresponding ONS mid-year estimates and official statistics births (respectively) to identify the degree to which there is denominator and numerator mismatch. Later in the chapter, fertility rates are calculated for different countries of birth using the carefully matched denominators and numerators selected.

With Chapters 4 and 5 having identified the samples of LS members where residence in England and Wales is accurately recorded, Chapter 6 turns more directly to evaluate the migration data in the LS. The key objective of this chapter is to identify any bias in registration of migrants with a National Health Service (NHS) General Practitioner (GP). Three questions are posed. Firstly, how good has the ONS LS been at collecting new LS migrant members through the GP registration process between 1991 and 2001 as opposed to 'collecting' them in 2001 when they appear in the census for the first time? Secondly, are women who are more likely to give birth more likely to register with a GP and enter the LS? Thirdly, do migrants to England and Wales register with a GP when they become pregnant and therefore make using the date of entry to the ONS LS / NHSCR inappropriate for considering the duration to birth? These are important considerations, as using the date of GP registration as a proxy for migration has the potential to introduce bias if exposure to risk in the dataset (registration) is a function of fertility.

Given the findings of Chapter 6 about the association between registration with a GP and subsequent birth, Chapter 7 identifies samples of LS migrants who entered the LS just before the 2001 census and comparison groups which entered at an earlier point in the 1991-2000 period. This analysis uses the findings in Chapter 6 to take an alternative route to selecting a sample for duration-based analysis with a measure of migration independent of NHSCR registration. The main aim of the analysis is to identify a sample of LS members arriving in England and Wales just before the 2001 census, to fully detail the background socioeconomic characteristics of the sample and finally, to estimate the fertility rates of this group with reference to other samples identified. The output from this chapter is a broader grouping of recent migrants as of the 2001 census and fertility rates for these groupings.

In Chapter 8 the key research question is returned to: what is the timing of fertility among recent migrants to England and Wales relative to non-migrants? This is answered by using a range of measures of the duration from arrival in England and Wales. A range of estimates of the true date of migration to England and Wales using the NHSCR data, along with other measures, are made and the groups identified in

Chapter 7 are used in a series of discrete-time hazards models to estimate fertility after the 2001 census. Based on the earlier findings related to attrition between 1991 and 2001, unobservable attrition (without 2011 census data linked in) is likely. Therefore, this analysis uses the first 24 months after the 2001 census. The benefits of using this time period include the ability to include socio-economic information from the 2001 census and certainty of the population exposed to risk at the time of the 2001 census, therefore minimising the potential bias which could be introduced via attrition from the ONS LS. This represents the final piece of analysis in this thesis and adds new insights on the timing of fertility among migrants to England and Wales, which has been a relatively under-researched area and not one for which the LS has been used.

The final part of this thesis, Chapter 9, concludes and contextualises this work fully. There is a full discussion of the findings in relation to the research questions in each chapter and the overall objective of identifying if there is an elevated level of fertility among recent migrants to England and Wales. The availability of 2011 census data in the revised version of the ONS LS in 2013 will provide scope for additional research on fertility in the 2001 – 2011 period with an added degree of certainty over the sample for analysis. In particular, the inclusion of a question at the 2011 census asking persons born overseas their year and month of migration to the UK will allow cross-checking with the NHSCR date of entry. This variable should allow the identification of the duration from arrival with greater ease and accuracy, rendering some of the analysis here unnecessary in the future. In a decade of high migration, like that of the 2001-2011 period, this is a valuable addition to the dataset.

Chapter 2

Empirical background – the fertility of recent migrants

Chapter abstract

In the late 1990s and early 2000s migration to England and Wales increased and since 2001 there has been a consistent rise in the United Kingdom total fertility rate (TFR) from a post-war low of 1.6 in 2001 to a high of 2.0 in 2010. Given the increase in fertility and the increase in migration observed, it is pertinent to consider associations between the two. Research by Toulemon (2004) and Andersson (2004) identified an association between migration and fertility after the migration event. Research on the fertility of migrants to England and Wales has focused on changes to the quantum of fertility and not the timing of childbearing. This chapter outlines the empirical background / context to this thesis and consists of a section exploring literature on the fertility of recent migrants and a section detailing past demographic research which has used the Office for National Statistics (ONS) Longitudinal Study (LS).

2.1 Introduction

This literature review explores research and theory on the fertility of migrants before outlining the substantive findings of previous demographic research which has used the Office for National Statistics (ONS) Longitudinal Study (LS) for England and Wales. Since 2001 the total fertility rate (TFR) for England and Wales has risen from a recent, post-war low of 1.6 in 2001 to just below replacement level at 2.0 in 2010. From the late 1990s there has been a significant increase in international migration to Britain, which accelerated after 2004. It is important to understand the relationship between these trends – as Sobotka (2010) suggests, migration and fertility are very much interrelated events. However, there are competing hypotheses on the association between migration and fertility and also the causes for associations, where they have been identified.

Increased migration can affect fertility in at least three ways. First, migrants tend to be of childbearing ages and therefore contribute to the stock of women at risk of giving birth and the overall number of births. In England and Wales the proportion of births to women born outside of the UK rose from 15% to 22% between 2001 and 2007 (Tromans et al., 2009). Tromans et al. (2008) identified a geographical match between local government areas where there has been a substantial increase in fertility and areas where there has been a high rate of immigration between 1986 and 2006. Second, some migrant groups tend to have family sizes larger than the host country and act to increase overall family size. In the British context we have seen significant assimilation of family sizes among many migrant groups such as Indian. However women born in Pakistan and Bangladesh continue to have higher completed family sizes than UK-born women (Dunnell, 2007) and are more likely to move for family formation (Peach, 2006). Third, period fertility rates can be inflated by a timing effect. Research from other countries has suggested that there is typically a short duration between migration and subsequent fertility (Andersson, 2004; Toulemon, 2004). As a result, the TFR for migrant groups, and hence the overall population, is biased upwards. In the UK, as will be shown, little research has been completed to examine whether there is an elevated level of fertility after the migration event.

Indeed, with reference to the timing of fertility around the migration event, three (Stephen and Bean, 1992) or sometimes four (Kulu, 2003), theories on the interrelation of migration and fertility have been proposed. Kulu (2003) identifies the *socialisation hypothesis*, the *adaptation hypothesis*, the *selection hypothesis* and the *disruption hypothesis*. The socialisation hypothesis has the premise that the fertility behaviour of the migrant relates to the fertility preferences of the childhood environment while the other hypotheses suggest modified fertility in relation to the migration process. The different hypotheses are fully outlined in the next section but the key aspect for this research is whether there is some form of disruption or modification to the fertility of the migrant because of the migration event. With an increase in migration flows to England and Wales from areas of the Global South, where there has been geopolitical instability, and also from the 2004 European Union accession countries, it is possible that there has been a corresponding increase in fertility and that this association has led to the increase in period fertility rates since 2001. In other European contexts an elevated level of fertility has been observed among migrants to the host country; disruption effects have been identified in France (Toulemon, 2004) and Sweden (Andersson, 2004) with migrants exhibiting a higher level of fertility following the migration event. The interaction of migration with subsequent fertility is an area where there are different theories and research findings which are applicable from country to country and time period to time period.

As briefly outlined, the current research background on the fertility of migrants to England and Wales has been more concerned with changes to the TFR (Tromans et al., 2009) and not the timing of childbearing relative to migration events. Changes to the TFR catch the overall quantum of fertility in England and Wales but would not identify any duration effect from migration. A sufficiently large stock of migrants of childbearing age who have a preference for waiting to give birth after migration could lead to an increased total fertility rate. Section 2.2 fully explains the theory and empirical findings on the fertility of recent migrants and looks at this with reference to migratory trends to England and Wales.

In section 2.3 the ONS LS is introduced and past research which has used the dataset explained. The ONS LS is an approximate one per cent sample of the population of England and Wales which, through the use of a link with the National Health

Service Central Register (NHSCR), records persons migrating to England and Wales for the first time, records embarkations (departures), deaths and new births. Through the use of four consistent and equidistant birth dates LS members are identified at each census and on the NHSCR. Numerous users of the ONS LS have researched fertility and the main objective of this part of the chapter is to give a firm overview of past work using the dataset.

Drawing on these preceding sections, section 2.4 makes conclusions on avenues for this present research by identifying that there has been limited research on the timing of fertility of migrants to England and Wales. To remedy this, it is proposed that the ONS LS is used to estimate the fertility of recent migrants to England and Wales relative to a comparator group of non-migrants. Immediate next steps to identify typologies of LS member and a sample from the LS which can be confidently used are outlined.

2.2 The fertility of recent migrants

This section of the chapter outlines recent migration and fertility trends in official statistics for England and Wales, discusses the theoretical considerations in understanding the fertility of recent migrants and explores empirical work which has sought to understand the association between migration and fertility. Overall, the section identifies that in England and Wales there has been insufficient research on the timing of fertility among recent migrants to the country.

2.2.1 Notable empirical work / theory on migration and subsequent fertility characteristics

As explained in the introduction to this chapter, there are three ways in which migration and fertility can interrelate, these being through changes in the stock of women at risk of giving birth, the differential family size preferences and finally through changes in the timing of fertility to migrants because of the migration event. While in the international literature there has been much interest in the duration from a migration event to subsequent birth, this is not something that has been reflected in research in the UK. Generally, research on immigrant fertility in England and Wales has been mainly concerned with the relative contribution of the group to the overall number of births or 'quantum' (Dunnell, 2007; Tromans et al., 2009). In part, at least, this is a reflection of the data available for researching the duration timing / trends for immigrants to England and Wales (Sigle-Rushton, 2008). Within the media, interest in the fertility of migrants has also been focused on the overall 'quantum' of fertility among migrants rather than the timing of such fertility (Boseley and Saner, 2009).

Theory on migration and subsequent fertility characteristics

As already suggested, there are competing views on the impact of migration on subsequent fertility. Indeed, this is a point identified by Kulu (2003), who provides an overview of many of the theoretical positions before analysing the situation in Estonia. Perhaps the reason for differing views and findings in research can be attributed to the differing contexts in which post-migration adjustment happens. Indeed, Kulu identifies "different views exist concerning the impact of a new social environment on childbearing preferences and behaviour of migrants" (Kulu, 2003 p.3). The new social environment is integral to subsequent fertility, according to Kulu.

It is possible to divide hypotheses on the fertility of migrants into four broad types according to Kulu (2003), although these are not necessarily mutually exclusive. These consist of the *socialisation hypothesis*, the *adaptation hypothesis*, the *selection hypothesis* and the *disruption hypothesis*. The socialisation hypothesis suggests that the fertility of the migrant is determined by the fertility characteristics of their childhood environment. All the other hypotheses suggest some form of modification in the fertility behaviour of migrants or a difference from the trends in the country of origin. The adaptation hypothesis suggests that the migrants' fertility will become like that of the society that they move to. Comparable with this, but suggesting that there is an unidentifiable similarity between the migrant and the society to which they move, is the selection hypothesis. This says that the fertility preferences of the migrant in the new location will be a better match that those of the country of origin. Meanwhile, the disruption hypothesis states that the fertility of the migrant will be low immediately after the migration event because of the moving process.

Other typologies have differed slightly: Stephen and Bean (1992) worked with just three types of fertility change in their study of Mexican immigrants to the United States of America (US), these being adaptation, assimilation and disruption. In effect, Stephen and Bean ignore the socialisation hypothesis which considers the childhood environment and cultural influences. Again, these are not necessarily mutually exclusive.

For this present enquiry into the changes in the TFR in England and Wales, it is the disruption hypothesis which is of key interest. If the migration event is a disruption in the life course of a migrant then it is possible that there could be some form of subsequent catch-up in fertility in the new country of residence for the migrant. Seminal works which have identified such a trend include Ford (1990) who identified an increase in immigrant fertility after migration to the US, Schrool (1990) who looked at immigrants to the Netherlands and Goldstein and Goldstein (1981) who looked at the fertility of migrants in Thailand. These studies identified some form of elevated fertility among the migrants after their migration event. Indeed, Coleman (1994), writing about fertility and intermarriage among immigrants at a macro-level, identified that overall, “fertility of the immigrant populations, measured as the TFR, enters the statistical series at an even higher level than that of the sending country” (Coleman, 1994 p.121-122). While this is not universally true, this does identify what the duration since migration may be doing to the period measure of the TFR – showing an elevated level because of fertility increases as a result of the migration event.

Given the increases in migration to England and Wales and the increases in fertility from 2001, there is a need to explore whether there is some form of relatively short-term increase in fertility among recent migrants. The next section explores recent empirical work in this area and discusses some of the findings on the relationship between migration and subsequent fertility.

Research on the timing of immigrant fertility

Among recent empirical works on the duration from migration to birth, Toulemon (2004) and Andersson (2004) both identified 'disruption' in the form of an elevated level of fertility in the period after the migration event. Toulemon (2004) identified that births to migrants in France often quickly follow migration. However, after an elevated fertility following the migration event Toulemon identified that there is a longer-run convergence to 'French' fertility. Meanwhile, Andersson (2004) looked at the situation in Sweden for a 30 year period and identified a higher level of childbearing shortly after migration.

The literature on the fertility of migrants can be split into that which is concerned with the timing of birth (the 'tempo') and that with the number of births (the 'quantum'). In the UK there has been a focus on the quantity of births to immigrant's *vis-à-vis* the fertility of the UK born population. However, research by Toulemon (2004) has identified that in France there is a distinct trend in the fertility of recent migrants where births often follow a migration event. Among women aged 25-30 years, upon arrival there was a marked trend in their fertility, with a very high fertility after arrival and then a drop to match 'French' fertility. In the context of recent research in England and Wales, the finding that using the TFR to look at migrant fertility over-estimates immigrant fertility is an important one. The approach taken by Toulemon is to identify the number of children ever born at the time of arrival and then calculate the fertility rates by the length of time since arrival. This perspective considers the whole childbearing history of the immigrant. Héran and Pison (2007) found that the fertility of foreign women is higher than that of the French women and that, although foreign women account for 12 percent of births in France, the effect of overall fertility is minimal. However, when Héran and Pison consider this in relation to the TFR they identify that the extra births increase the TFR by just 0.1. In sum, they conclude that the higher fertility of French women is of more interest. While this is true to an extent, the short-term impacts of migration on fertility do merit research and when there have been changes in the level of migration to a country, as observed in England and Wales, then this merits further research.

Taking a similar approach to Toulemon, Andersson (2004) considered the fertility of migrants to Sweden. Data from the Swedish population register was used to look at the whole period from the 1960s to the 1990s using event-history techniques. A higher level of childbearing shortly after immigration was identified, the nature of the relationship leading Andersson to argue that migration and family-building are inter-related processes and the time since migration must be considered when the fertility of migrants is studied. Overall, three factors are considered – the period effects on fertility, effects of country of origin and effects of migration on childbearing behaviour. Looking at each of these in turn, the risk of childbearing for childless women aged 16-28 years was highest for non-Nordic women, for second births the rates for Swedish women were greater and for third births the rates decreased at a much slower pace for non-Nordic women. For all immigrant childless women (regardless of country of origin) there was a higher propensity to become mothers than childless Swedish women at corresponding ages and calendar year periods. With regard to effects of migration on childbearing for immigrants, it was found that immigrants to Sweden have first birth rates in their first two years of residence which are 100 percent higher than those of Swedish women. In addition to these three specific findings, it is identified that migration seems not to have any disruption on childbearing among immigrants.

Both Toulemon (2004) and Andersson (2004) have therefore looked at the timing of immigrant fertility at each birth parity rather than the absolute number of births. Period indices of fertility, the ASFR or TFR, are more affected by changes in the timing of childbearing which are intrinsically related to the age of migrants entering a country and their childbearing patterns. This means that an approach taking full account of the changing timing of childbearing among the population is more fruitful. Where there is an increase in the number of immigrants of a childbearing age and there is a high birth rate to these women, then the overall fertility rate for the group will be higher and the ASFR and TFR will show a change as a result.

Taking a European perspective, Sobotka (2008) identifies the growing role of migration as a driver of population growth and that migration has a multi-faceted impact on childbearing trends and population change. This paper gives an extremely strong overview of the literature, theory and thinking on the fertility of migrants. Period measures of fertility are discussed early in this work, as migration has a substantial impact on these rates because of "the interrelation between the events of migration and fertility." (Sobotka, 2008 p.228). Linking to the findings of Toulemon (2004) and Andersson (2004), Sobotka states that "migrant women typically retain substantially higher levels of period fertility than the 'native' populations, but this difference typically diminishes over time and with the duration of their stay in a country." (Sobotka, 2008 p.225). This is in common with the findings of other work in this area and, with reference to the plots of ASFRs by age of migrant at arrival by Toulemon (2004), this is certainly true for France. However, what Sobotka identifies is compatible with more than one social process as defined in migrant fertility theory and does not distinguish between an assimilation effect or a consequence of the association of the timing of migration and fertility.

The work of Toulemon (2004) and Andersson (2004) is not the only research to have looked at the fertility of migrants, but these papers are where the identification of the 'migrant effect' has been clearest. Mussino et al. (2009) identified that the fertility rate among North African immigrants in Italy was twice that of central and eastern European mothers. This used a specially constructed demographic dataset and studies the 2002-2006 period. It is identified that citizenship is important in explaining the high heterogeneity in the reproductive behaviour among the mothers. A limitation of this work is that it is only looking at second births, but the findings made are still relevant for the current study to consider. A later paper by Mussino and Strozza (2012) looks more directly at the relationship between the timing of a first birth in relation to migration. This work is again for Italy and is similar to that of

Toulemon (2004) and Andersson (2004). A cohort of migrants from Albania, Morocco and Romania, the three largest migrant groups at the current time, arriving in 2003 are studied. In the first 12-18 months of residence there is an elevated risk of birth and so a 'strong arrival effect' is identified. But beyond this elevated risk of fertility immediately after arrival a strong interrelation between migration and family behaviour is identified.

2.2.2 Research on migration and fertility in England and Wales

As already explained, in England and Wales there has been substantial interest in the number of births to immigrants. This interest has been provoked by the increasing number of immigrants to England and Wales and the rise in the total fertility rate from a low in 2001. The focus of the research has been on the 'quantum' or number of births and not on the timing of births. However, by focusing on the number alone, the actual timing impacts which migration can have on fertility rates are missed.

One of the key considerations with estimating the fertility of recent migrants to England and Wales is suitable data. Although ONS vital statistics give the country of birth of the mother, there is no corresponding denominator (apart for census years). This is a point identified by Sigle-Rushton (2008) who uses the Labour Force Survey (LFS) to look at the contribution of migrants to the recent increase in fertility observed since 2001. "Although TFRs for UK-born women have increased, high levels of migration and a higher share of new immigrants as a proportion of the foreign-born have contributed to the increase in the TFR between 2001 and 2006." (Sigle-Rushton, 2008 p.473). This suggests that there could have been a substantial contribution by migrants to the increase in the England and Wales TFR. Based on other research (e.g. Tromans et al. 2009) this seems plausible, yet this research is again concerned more with the quantum or level of fertility rather than the timing.

A crucial point made by Sigle-Rushton (2008) is that, as migration is generally recorded badly, the denominator (number of women resident) might not be accurate and therefore could be inflating the rate. This is an issue which may be addressed by using the ONS LS, providing the denominator can be suitably understood. Through using the NHSCR and entries via the census, the ONS LS collects a large sample of migrants and links births which are registered with a GP. This avoids respondent bias (Blackwell et al., 2005).

Research from the ONS has looked at the fertility increase since 2001 and the drivers of this. An increasing share of fertility from women born outside the UK has been identified. "In 2001, 15 percent of births in the UK were to mothers born outside the UK and by 2006 this had increased to almost 21 percent of births in the UK" (Dunnell, 2007 p.19). In particular, it was identified that the fertility of specific age groups has increased more than others – in the 25-29 year and 30-34 year age groups the fertility of women born outside the UK increased more than for women born in the UK. However, consistent with more recent research on the number of births to overseas born women (Tromans et al., 2009), it is identified that increasing fertility among UK born women has been important in the increase in fertility since 2001. "The Total Fertility Rate for the UK has increased from 1.6 children per woman in 2001 to 1.8 children per woman in 2006, the highest level since 1980. In England and Wales, the estimated total fertility rate for UK born women has risen from 1.5 to 1.7 since 2002, while for women born outside of the UK the estimated rate rose from 2.3 to 2.5" (ONS, 2007, p.1).
The merits of such rates would be questioned by Sobotka (2008) who identifies the relation between migration and fertility as impacting on period measures of fertility. The work of Héran and Pison (2007) would suggest that there is not such a dramatic impact on the overall TFR from the fertility quantum of recent migrants.

Research on births to foreign-born women (Tromans et al., 2009) and the degree to which there has been consistency in fertility trends across England and Wales (Tromans et al., 2008) shows the importance of immigrants to trends in fertility since 2001. Again, both of these papers are more concerned with the increasing number of births rather than any impacts from changes in the timing or the duration from migration to birth. However, the number of births is, by definition, related to the number of immigrants and the timeframe for which they have been resident. In Tromans et al. (2008) the timeframe for analysis is 1986-2006, which encompasses both the decline and rise in fertility observed. Local Authority areas where there has been a substantial increase in fertility between the two dates are identified and in many cases these are areas where there has been a high rate of immigration in the intervening period. For example, Barking and Dagenham, Newham and Rutland make up the top three areas where there is the highest TFR in 2006. Peterborough and Boston, two areas which have attracted large numbers of accession eight ('A8') migrants, also feature in the top ten. This is in line with findings related to the labour market preferences of 'A8' migrants from the EU expansion in 2004 (Coombes et al., 2007). In the case of Boston "much of the increase is due to a rise in the number of births to mothers born in eight out of the ten countries that acceded to the European Union in 2004." (Tromans et al., 2008 p.18). Within the paper it is noted that TFR for immigrants could be inflated by timing effects and that there is the potential for a numerator and denominator mismatch in areas where there has been high migration, particularly at small geographic scales.

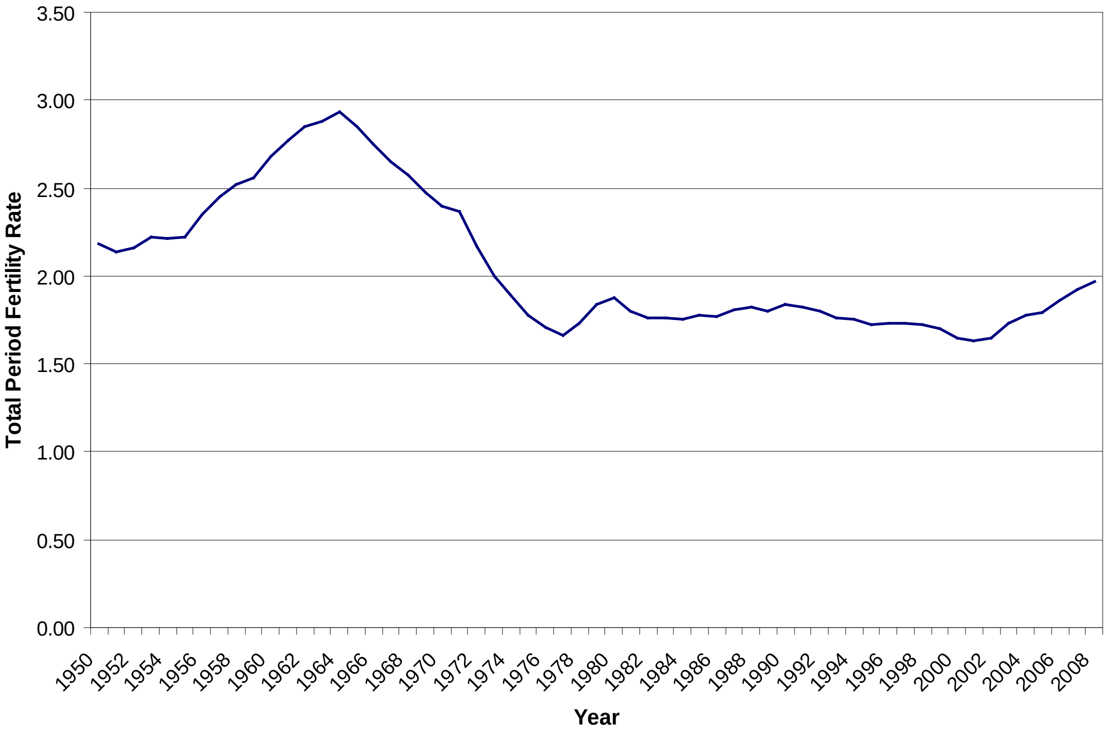
In Tromans et al. (2009) it was identified that births to foreign-born women accounted for around two-thirds of the fertility increase between 2001 and 2007. An important point with regard to the UK population made early in the paper is that the number of UK-born women in childbearing years fell in the 2001-2007 period, while there was an increase in the number of foreign-born women aged 20-24 years. Overall, despite a great part of the increase in births being among foreign-born women, the overall TFR increase was mainly due to increasing fertility among UK-born women. This is a reflection of the numbers of UK-born women relative to the foreign-born population.

2.2.3 What do official statistics for England and Wales show?

Increasing fertility

It is helpful to consider what the official statistics from the ONS tell us about changes to the fertility rate in England and Wales. Figure 2.1 shows the changing total period fertility rate (TPFR) for England and Wales in the years 1950-2008. The most recent rise in fertility from 2001 to 2008 is relatively small compared with the 'baby boom' in the 1950s to the mid-1960s. However, in the context of the prolonged period of below-replacement fertility since the early 1970s it is a notable rise, and one from a low base. In 2001 the total fertility rate (TFR) was just 1.63, which is among the lowest recorded rates in the post-war period. In contrast, the rate in 2010 (2.0) is the highest since the transition to below-replacement fertility in the early 1970s.

Figure 2.1: Total Period Fertility Rates (TPFR) for England and Wales, 1951-2008



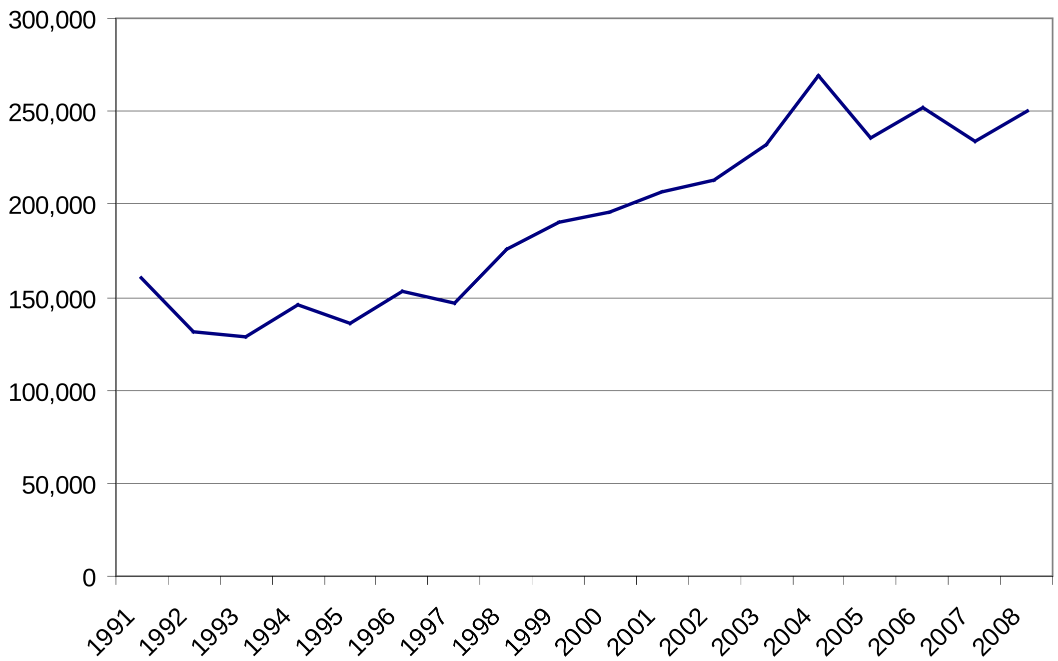
Own elaboration based on ONS; Social Trends and Population Trends, Accessed 25-05-2010.

Increasing migration

At the same time as the increase in fertility, there has been an increase in migration to England and Wales. Migration has increased because of geopolitical instability in the Global South (i.e. Iraq, Democratic Republic of the Congo) and as a result of the expansion of the European Union in May 2004 to include eight former Soviet bloc countries (the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, and Slovenia). Nationals from those countries wishing to work in the UK from May 2004 could do so by registering under the 'Worker Registration Scheme'. Figure 2.2

shows data from the ONS on female 'long-term' international migration to England and Wales in the 1991-2008 period. The graph shows that, in the main, the 1990s saw a consistent level of migration, with an annual inflow of around 150,000 females a year. From the latter part of the decade, in particular 1998 and 1999, there was an upwards turn in the number of female migrants. Through to 2003 the annual numbers of migrants remained constant, but in 2004 there was an upwards kink to 269,000 migrants. After this peak the annual inflow has fluctuated around 250,000 female migrants per annum.

Figure 2.2: Long-term female migrant inflow to England and Wales, 1991-2008

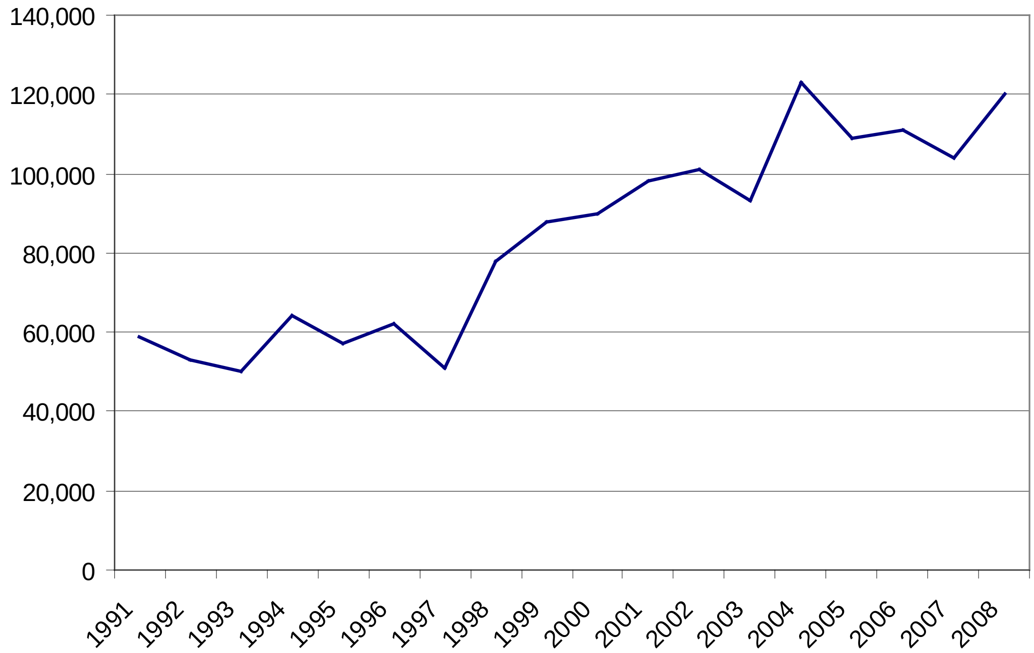


Own elaboration based on ONS; Social Trends and Population Trends, Accessed 25-05-2010.

Figure 2.3 shows the same data but only for the 25-44 years age group, which encompasses the main childbearing ages. There is a noticeable increase in migration from 1997 to 2004. The upwards turn in the number of migrants is distinct. Neither figures show a consistent trend for the 2004-2008 period. In 2005 there was a small decrease in the number of migrants to fewer than 250,000 per annum; this was also seen in 2007. These may be random fluctuations arising because estimates are based on the International Passenger Survey which is a random survey. In 2006 and 2008 there were over 250,000 female migrants per annum entering England and Wales.

Depending on the time period, the composition of migrant flows to England and Wales (sending countries) varies greatly. In the late 1990s drivers of migration included the geopolitical instability in the Sierra Leone and Democratic Republic of the Congo.

Figure 2.3: Long-term female migrant inflow (aged 25-44) to England and Wales, 1991-2008



Own elaboration based on ONS; Social Trends and Population Trends, Accessed 25-05-2010.

Increased fertility because of increased migration?

In the annual National Statistician's article, Matheson (2009) unpicks some of the trends in the migration from the A8 countries and analyses data on the characteristics of migrants – the bulk of immigrants are young adults of working age. The most recent migration figures show that there has been an increase in the outflow of A8 migrants from the UK; the timeframe within which this has occurred coincides with the contraction in the UK economy. This same paper outlines some of the 'headline' statistics with regard to the fertility of the A8 migrants. While there was a seven-fold increase in their births between 2004 and 2008, just 3.2 percent of all births were from this group.

In literature on the fertility of migrants, it is accepted that migrants have a higher rate of fertility (Coleman, 2007; Sobotka, 2008). The timing of births to migrants to England and Wales is, as detailed in the next section, a relatively under-researched area.

2.2.4 Conclusions / summary

Empirical work on the fertility of recent migrants in other European countries shows what looks like a consensus: migrants do have higher fertility than the populations of the country to which they move. The work of Toulemon (2004) and Andersson (2004) is among the strongest in terms of the findings made and the robustness of the approach taken. In particular, as shown in this section, there is a difference between migrants showing a higher level of fertility, as measured in period fertility measures like the TFR, and a duration effect whereby there is a trend in fertility associated with the duration from migration to subsequent birth. Research on the fertility of migrants to England and Wales has focused on the number of births and the fertility rate rather than timing. Yet, as has been outlined here and recognised in the literature, it is possible there is a timing impact on fertility when migration increases.

To estimate any timing effect for recent migrants to England and Wales, it is important to use the correct data. This is something identified in Kulu (2003) and Sigle-Rushton (2008). Kulu (2003) also notes that in the main, cross-sectional data has been used and that there has been a limited use of longitudinal data. "Longitudinal data have found only limited use, despite their dominant position in many areas of population research" (Kulu, 2003 p.9). The availability of data which is suited to researching the fertility of migrants is problematic. Among the datasets on offer in the UK the ONS LS is one of the best for understanding the duration of residence and fertility of LS members. The large sample of migrants which have been collected in the dataset, and the way in which births to LS members are linked from vital registration data, mean that use of this should make analysis of immigrants by country of birth possible and allow estimation of the duration from migration to birth. The next section explores past demographic research using the ONS LS.

2.3 Empirical fertility research using the Office for National Statistics Longitudinal Study

This section of the chapter outlines past research which has used the ONS LS, the approach taken and the substantive findings made. There is an extensive catalogue of demographic research which has used the LS, some from researchers at the ONS. Fertility research using the LS has the benefit of almost complete linkage of births from the vital registration system (see the next Chapter for a full explanation of the linkage of births to LS members) avoiding reporting errors as identified in some retrospective fertility histories (Ní Bhrolchaín et al., 2011; Murphy, 2009) and a large sample size (over a million all-time members). Although a wide range of demographic research has made use of the LS, this review is primarily concerned with the way in which fertility and migration research has used the LS and reports on this type of research. Several different research areas within which the LS has been used are outlined.

2.3.1 Use of the ONS LS to estimate the interrelation of family policy and changing fertility

Among the most well-known fertility papers to have used the LS is that by Ekert-Jaffé et al. (2002) and Rendall et al. (2009), which both looked at family policy regimes and fertility comparing England and Wales with France. One enabling factor for such research is comparable French data, the Permanent Demographic Sample (known as the Échantillon Demographic Permanent: EDP). Ekert-Jaffé et al. (2002) compare the socio-economic circumstances of births and their timing in England and Wales with France. The paper seeks to explain the contribution of French family policy to the fertility characteristics in France relative to England and Wales. It is suggested that the revisions made to the French policy in the 1980s and 1990s led to an approach which is more 'Nordic' in its style – where there has become a greater emphasis on enabling women to combine careers and motherhood, should they wish to do so. "In France, the policy of providing childcare assistance appears to allow better-educated women to become mothers sooner than in England" (Ekert-Jaffé et al., 2002 p.491). Policy differences seem to be impacting on different socio-economic trends in fertility between the UK and France. Analysis of parity progression by Ekert-Jaffé et al. shows that in the UK there are stronger associations with occupation and particularly withdrawal from the labour market.

Meanwhile, Rendall et al. (2009) extends this work to study the distribution of fertility and identify if 'universalistic' regimes, like that in France, allow an easier reconciliation of work and fertility than 'means-tested' or 'conservative' welfare regimes. The paper approaches this by disaggregating fertility by women's age, parity and pre-childbearing occupation. Therefore this approach, it is argued, allows the identification of how the policy regime may in some way be affecting the distribution of fertility even when the overall level is not affected.

2.3.2 Cross-national comparative work using the ONS LS

The composition of the LS means that it is comparable with other datasets across Europe. Rendall in particular has made use of the LS in this way for a series of research papers in the 2003-05 period. These have included work with the French EDP, LS and population registers from Scandinavia. Rendall and Smallwood (2003) used the ONS LS to look at higher qualifications and their association with first-birth timing and further childbearing in England and Wales. Data from the ONS LS is drawn for a cohort of women born in England and Wales between 1954 and 1958. There are some key findings for socio-economic and fertility inter-relations – average age of entry to motherhood is five years later among women with higher qualifications compared to those without. Increasing age of motherhood is always associated with a lower likelihood of going on to have another child. However, the paper identifies that this decline is less pronounced for women with a higher qualification. The paper also identifies tighter spacing between births for women with higher qualifications.

Inter-generational teenage fertility is a concept which has received much interest in demography. Because of the sample of the LS, and the way information for non-LS member household information is linked from the census, it is possible to study the fertility of other family members. In Rendall (2003) a comparison of teenage fertility is made with France using the EDP. The paper identifies that in England and Wales mother-daughter repetition accounts for only a minor part of the total difference in teenage childbearing between the two countries. In making this finding, a valuable contribution is made to persistent questions over the degree to which the higher level of teenage fertility in England and Wales is because of a repetition effect among the children of teenage parents. Because teenage pregnancies comprise a relatively small sample of the overall number of births, the ONS LS is probably the only dataset in England and Wales which would allow this type of research to be completed.

In Rendall et al. (2005) first birth trends by age and education are estimated using the LS, EDP and Norwegian Population Register. A higher degree of association between terminal education level and age at first birth in Britain is identified. Using cohorts from the 1950s and 1960s, it is identified that in France and Norway the peak ages for risk of first birth childbearing shifted more uniformly across education levels for the two cohorts. One of the reasons for the importance of such research is that increasing education levels among women are associated with the shift in childbearing to later ages. This later starting of fertility in turn has contributed to more constrained fertility and a lower level of overall fertility (Berrington, 2004). Because of 'uncertainties' about the date of arrival among migrant women, only native-born women are included in the analysis. The findings of this paper relate to some of the other comparative research using the EDP and LS, including that by Ekert-Jaffé et al. (2002) which identified that the level of socio-economic difference in the level of fertility between France and Britain had actually decreased.

2.3.3 Data functioning and quality

There have been relatively few papers using the LS to examine how reliable the dataset is, and how it may be used for research. Among those that have is that by Babb and Hattersley (1992) with Hattersley having also written on the functioning of the LS and its data quality (Hattersley and Creeser, 1995). This specific paper (Babb and Hattersley, 1992) used the ONS LS to estimate age-specific fertility rates for the 1971-1988 period using a cohort of women born after 1950. It is suggested that the fertility data from the LS compares well with the figures for England and Wales (see Chapter 3 for example graphs from this paper), although what are identified as being quite small gaps between the LS and official England and Wales figures are actually sizeable. One of the major benefits of the data at this point in time was the ability to derive fertility rates for married and unmarried women. Indeed, one of the original purposes of the LS was to estimate non-marital fertility and estimate birth spacing. Babb and Bethune (1995) used the ONS LS to look at extra-marital fertility. Again, the large sample available for analysis and the way in which the LS collects birth information for births to sample members is an important factor in choosing to use the LS. At the inception of the ONS LS there was no data available on non-marital fertility – this was not recorded. As a tool for capturing the change in this over time, the LS is more or less unique in the sample of members which can be used for analysis.

To the present day, the best volume on the functioning of the dataset remains Hattersley and Creeser (1995) which details the history, organisation and quality of the data. This is now somewhat dated given the way in which the processing of the data has changed since the early 1990s and the linkage of the 2001 census data has taken place (with 2011 data being linked currently). Chapter 3 discusses the way in which the LS draws together different sources of data and discusses data quality considerations.

2.3.4 Migration research using the LS

Among user guides on the LS, Hattersley (1999) covers the international migration data in the LS and the way in which this can be best utilised by researchers. Importantly, this guide identifies that "The quality of migration data in the LS is difficult to measure" (Hattersley, 1999 p.5). This seems to be due to the lack of reliable migration data for England and Wales against which the data in the LS can be benchmarked. It is suggested in the report that the processing of the migration data at the NHSCR means that migration events as recorded in the LS are actually identified long after they have occurred. Therefore, the entry date on the NHSCR as recorded in the LS can be inaccurate. Through comparisions with International Passenger Survey (IPS) data (the main dataset used for measuring migration from and to England and Wales by the ONS), it is identified that immigrants are over-represented in the LS while emigrants are under-represented. It is suggested that where there is lost-to-follow-up at the next census then it can be assumed that the person left just before the census, however, the rationale for this suggestion is not fully given. In summary, the report identifies that the LS is better at capturing migrants to England and Wales rather than recording emigrants. This seems to be because of the way in which new entrants to the country are recorded with an NHSCR GP registration, but departures are not always recorded as an LS member does not have to deregister with a GP when they leave. For research such as this, which is interested in using the migration data in the LS, these are considerations which will be explored further in the next chapter.

A more recent paper which makes use of the patient register data (NHSCR data) in the LS is that by Smallwood and Lynch (2010). This uses the 2001 census data and patient register data in the ONS LS to examine potential sources of difference in usual place of residence at the 2001 census. The work finds that 96% of LS members enumerated at the census resided in the same area as recorded in NHSCR data. The rationale for this work was that an examination of the way in which the patient register data and the census data compare had not been completed before and this gives the opportunity to see where the census placed people in relation to the patient register data. Although not referred to, this is particularly interesting work if there will not be a 2021 census – by making such a comparison the findings from this work give an overview of where the entire NHSCR could be used as the basis of a system for estimating the population annually. Among women who were registered at a different address on the NHSCR compared to that recorded on the census form (and the LS) most had moved GP (therefore changing the NHSCR record) within 6 months of the census. There was a trend whereby women re-appeared more quickly than men. In the conclusions to this work it is suggested that there is scope for greater use of the address 12 months before, as recorded on the 2001 census form, to look at trends in migration and updating of the NHSCR. Overall, this gives a great deal more information on the trends among migration in the period before the census and migration data in the NHSCR and LS than has previously been the case.

After a period of extensive research, one of the only papers to have looked at migration and fertility using the LS is that by Grundy (1986). This studied relationships between migration, housing tenure and fertility in the 1971-1981 period. Headline findings include: more tenants moved between the 1971 census and the first subsequent birth than owner-occupiers and the association between tenure and moving was more consistent than the relationship between moving and the husband's social class. Of perhaps more interest to this work were the findings in relation to long distance moves. A postponement of the first or second child was identified, although it was suggested that both longer distance migration and fertility behaviour are associated with other characteristics such as education. Important points to note with this work are that this is only concerned with internal migration within England and Wales and the period under investigation, the first decade of the LS, is one in which there was a lower level of linkage of births into the LS. This does, however, start to show the type of analysis which can be completed using the LS because of the long timeframe for analysis and the way in which the socio-economic details of individuals are recorded at each census.

2.3.5 Fertility timing and spacing work

Among the original reasons for starting the LS was to better understand birth spacing and parity progression ratios among the population of England and Wales. Werner (1988) used the LS to look at the spacing of births among women born in 1939-59. Relating back to some of the previous work which has been discussed here, this identifies that the first birth is crucial in determining the subsequent fertility for women. This work identifies that births inside and outside marriage are linked to childbearing histories of women in the sample – an advantage of the LS compared to other datasets for this time period.

A more recent work to have used the LS to look at fertility and more specifically, childlessness is that by Portanti and Whitworth (2009). This used a sample from the ONS LS which was continuously resident without any form of embarkation from the LS. Female LS members from the 1956-61 birth cohort were selected for analysis and the main aim of the work was to identify the degree to which the socio-economic characteristics of the women and their partners relate to childlessness. Although the article correctly identifies that in the main it is cohort studies which have been used for such research, the article does not discuss linkage of births to LS members in the LS and systematic trends in the longitudinal linkage of LS members between one census and the next, especially given the use of such a long time period. The main finding is that partnership status is key to determining fertility outcomes. Irrespective of this, women's socio-economic characteristics are associated with childlessness.

Another paper to have looked at social class relations with timing of first birth is that by Buxton et al. (2005). This paper, in a similar way to Rendall (2003), draws upon the information on LS members' parents to look at the influence of socio-economic characteristics of LS members' parents on the occupation, education and family building patterns of adults. It is identified that these patterns vary considerably by parental social class. This work, like that of Portanti and Whitworth (2009), does not discuss systematic trends in the longitudinal linkage of LS members between census dates. However, the descriptive results show interesting occupational, educational and family-building patterns of middle-aged adults and how these vary with parental social class.

2.3.6 Conclusions / summary

This part of the chapter has considered the ways in which substantive demographic research using the LS has approached the dataset and findings made from its use. There is strong body of research which has used the ONS LS over the years. Provided a sample of migrants can be accurately selected from the LS, it seems that this dataset may be suitable for estimating the fertility of migrants, given the way in which migrant details are recorded through the connection with the NHSCR, the 2001 census information and the way in which the 2001 census can provide a complete socio-economic background to the migrant.

Little recent research has used migration information which is recorded in the LS and the detailed fertility histories to look at their interrelations. Both the work of Grundy (1986) and Smallwood and Lynch (2010) identified the potential for the LS to be used to look at the fertility of migrants using some of the different forms of migration information which are recorded in the LS, both through the linkage of the NHSCR information to the LS members and also from the questions asked on usual place of residence 12 months before the census. Within the technical work to have looked at the migration information in the LS, the paper from Hattersley (1999) stands out. This identified that the LS is better at collecting information on when persons enter the country than when they leave. In the next section overall conclusions are made based on the discussion of the literature.

2.4 Conclusions

This literature review has identified that there has been a limited examination of the trends in fertility among recent migrants to England and Wales and whether there is some form of disruption or elevated level of fertility among migrants to England and Wales. In section 2.2 it was explained that among several theories on fertility and migration, the disruption hypothesis fits with the demographic research which is relevant for this present research to consider. The work of Toulemon (2004) and Andersson (2004) is of particular interest as this has looked directly at the timing of fertility among migrants and found that there is an elevated level of fertility in the period immediately after migration. In the UK, increasing migration and increasing fertility levels in the 1991-2010 period were shown to have occurred at the same time and it has been highlighted that past research has mainly looked at the quantum of the fertility of immigrants rather than the timing. While the research on this and the

TFRs calculated (Tromans et al., 2009) is extremely interesting; there has been little research looking at the timing of fertility among migrants to England and Wales.

In section 2.3 an outline of work which has used the ONS LS for fertility and migration research was provided. This has shown the wide applicability of the dataset for different forms of research and in a cross-national comparative context. In the conclusions to this section it was suggested that there is currently a gap within the body of work which has used the LS because there has been little work looking at migrants to England and Wales and their subsequent fertility. Provided the way in which the ONS LS operates and any systematic trends in the operation of the dataset (and the implications of these) can be considered, it is possible to use the ONS LS to estimate the fertility of recent migrants to England and Wales. Therefore, Chapter 3 focuses in detail on the way in which the LS functions and how data from various administrative sources and the census is combined to form the LS. There are various routine publications which the ONS produces on the quality of the data, and the implications of these are discussed. In addition to this, there is a discussion of past research which has used the LS and how the findings on data quality have shaped the research approaches adopted.

Chapter 3

Data – the functioning of the Office for National Statistics Longitudinal Study and past research approaches

Chapter abstract

The Office for National Statistics (ONS) Longitudinal Study (LS) for England and Wales is an extremely complex dataset composed of ONS census data, ONS vital statistics data and records from the National Health Service Central Register (NHSCR). The objectives of this chapter are to explain the rationale for the development of the LS; the sources of data used; the ways in which these are combined; to provide detail on the quality of the dataset according to secondary sources; and to review fertility research approaches to selecting a sample for analysis. Given the large sample size and the ways in which event information is recorded in the LS, it is possible to use the data for demographic and event history analysis, provided the operation of the dataset is understood and the population which is exposed to risk of birth can be identified.

3.1 Introduction

The Office for National Statistics (ONS) Longitudinal Study (LS) is a one per cent sample of the population of England and Wales, drawn from the 1971 census by using four consistent and equidistant dates of birth. ONS and National Health Service (NHS) records of 'vital events' are linked to members of the LS. Since the inception of the LS in 1974, new LS members have been added through the recording of new births on the four dates, the recording of migrants to England and Wales with an LS date of birth who have registered with an NHS General Practitioner (GP) and persons resident at a decennial census with an LS date of birth. Census data for existing and new LS members has been included from the 1981, 1991 and 2001 censuses. Decennial census data for individuals with one of the four LS birth dates is extracted from census files and added to the record for each LS member, where identified. The availability of census data collected for each LS member gives a range of socio-economic variables for him or her, but an LS member must appear on a census form for this information to be recorded. For the United Kingdom (UK) as a whole there are three Longitudinal Studies: one for England and Wales, one for Scotland and one for Northern Ireland. This research refers only to the England and Wales dataset, which is maintained by the ONS LS Development Team and made available to the research community through the Centre for Longitudinal Study Information and User Support (CeLSIUS).

This chapter is a first step in outlining LS data quality in relation to the profile of women in the LS and the accurate capturing of births to LS members. The priority in this chapter is explaining the way in which the LS functions and the approach of past research to the sampling quality of the LS. Section 3.2 presents a detailed explanation of the functioning of the LS, including the forms of entry to, and exit from, the LS and hypotheses concerning potential sources of error in the dataset. This leads to section 3.3, which outlines the data that composes the LS, processes used in the creation of the LS and a summary of reports by the ONS on data quality. Section 3.4 is concerned with the way in which fertility research has used the LS and identifies that, in general, previous research work using the LS in fertility analysis has steered away from explicit and open discussions of some data quality issues which are specific to a longitudinal and linked dataset of this type. In fertility research using the LS, there has not always been an acknowledgement that there are, as would be anticipated, different fertility rates in the LS compared with vital statistics data, nor discussions of the ONS publications on data quality. Section 3.5 provides conclusions on the operation of the LS and provides a summary of the main sub-groups of the population affected by representational problems in the LS.

This chapter is concerned with fully understanding the functioning and quality of LS data. For use of the ONS LS to estimate the fertility of migrants to England and Wales, the operation and quality aspects of the LS must be appreciated. The finding of this chapter is that the tendency in the body of demographic research using the LS to ignore or not report data quality aspects and sample selection criteria is disappointing, given that the LS is better at capturing some members of the population than others. This chapter is an important foundation for understanding the selection of a sample from the LS for research on the fertility of migrants.

3.2 What is the ONS Longitudinal Study and how does it function?

This section provides a detailed outline of how the ONS LS was developed and the data used in its composition. In the first sub-section there is an outline of the rationale behind the initial development of a longitudinal study for England and Wales and a short explanation of the development and evolution of the study. Sources of data used in the LS are introduced in this sub-section. The remainder of the section explains entry and exit points used in the LS and the degree to which these capture the population born on an LS birth date. The conclusions outline the main opportunities for attrition from the LS and key points on the operation of the LS.

3.2.1 The origins and initial purpose of the Longitudinal Study for England and Wales

The ONS LS was originally developed in 1974 using a sample drawn from the 1971 census. Through the use of four, equidistant birth dates, a one per cent sample of the population of England and Wales was selected from the 1971 census. A one per cent sample was achieved through the four dates as dividing 4 by 365.25 gives 1.09%. This is the target sample fraction for the LS to achieve in order for it accurately to represent the population of England and Wales. At the inception of the LS approximately 530,000 members were included in the dataset and since then (where possible) existing members have been linked at the decennial census to provide a longitudinal picture of their lives. New members have been added to the dataset, based on the same birth dates. New members enter the LS through migrating to England and Wales and providing an LS birth date when registering with a GP, being born on an LS birth date and being enumerated at a census with an LS birth date. This method of sample selection ensures that the dataset is of a consistently large size over time and that the population traced on the National Health Service Central Register (NHSCR) as a whole should be of a robust size for most research. This implicitly leads to assumptions that the LS is 'representative' simply by virtue of its large size, when this is not necessarily the case. The 1.09% sampling fraction should be true across all sub-populations, whether defined by socio-economic, birth cohort or other characteristics. Key to the functioning of the LS, and the main distinguishing feature of the LS relative to other datasets, is the way in which it is connected to the NHSCR. The NHSCR holds health record information for residents of England and Wales registered with a General Practitioner (GP) and where possible this is linked to the LS member. Individuals present at a census date with an LS birth date are identified on the NHSCR through a 'tracing' process. An LS member is traced when their corresponding record has been found on the NHSCR. The tracing of an LS member on the NHSCR is crucial for the ongoing maintenance of their record - if they cannot be located then it is not possible to update the LS record with event information from the NHSCR. Matching of census and NHSCR records after the 2001 census was completed by using the following matching criteria:

- Name (used for processing purposes only, not returned to ONS)
- date of birth
- sex
- postcode
- postcode one year ago
- person type (private/ communal establishment).

(Office for National Statistics, no date a).

Advantages of a longitudinal approach focused around birth cohort characteristics, rather than a cross-sectional perspective, were identified as long ago as Farr (1839). At the most detailed levels of statistical analysis, the use of retrospective questioning and cross-sectional analysis is insufficient. Indeed, writing in 1976, the then Chief Medical Statistician for the Office of Population Censuses and Surveys (OPCS) provides a rationale for the development of the LS as being because "of a considerable increase in the demand for more information about the population: more facts, in more detail" (Adelstein, 1976 p.2). To identify and link the individuals that were to be part of the LS, additional information at the recording of 'Vital Statistics' events was needed. To facilitate the LS, in 1969 an Act of Parliament added questions to the record of birth and record of death. On the record of birth, extra questions were added on the place and date of birth of the mother, and on the death certificate, questions added on the place, date of birth and maiden surname (of the deceased). The addition of these fields to the official recording of these events was crucial for the development of the LS (Adelstein, 1976).

As might be anticipated, there were ethical concerns when the development of the LS was granted approval. The primary concern was that 'dossiers' on individuals might be created (Stevenson, 1973). Apart from this, there is little literature on ethical concerns at the inception of the LS, or in the years since. Among those publications from the inception of the LS is that by the OPCS, written in 1973, which outlines the method to be used in the LS and the ways in which individual confidentiality will be ensured (OPCS, 1973). There are several ethical arguments made. Firstly, it is mentioned that the data being used in the composition of the LS are already held by the OPCS and therefore no external data providers or processors will be involved in the process. Secondly, it is argued that the process being used is solely statistical and not administrative, with serial numbers used for the linkage process. The last argument is that the divisions processing and managing the study will remain separate, which is important in undermining concerns about 'dossiers'. The OPCS report aims to allay concerns about the LS being used as a population register, composed of full names and information on individuals. "With such a small sample it is hardly possible to sustain the notion that the study records could be used as a reference source" (Office of Population Census and Surveys, 1973 p.4). Reports from the inception of the LS do not discuss the possible identification of, and disclosure of, information on individuals. The ONS LS Research Board (RB) must approve research before access to the LS is granted.

OPCS (1973) also outlines some compelling arguments for a longitudinal study. The use of the LS for research on occupational mortality and illness, and changing fertility (particularly spacing of births) were among the main motivations for the construction of the LS. With a longitudinal study it was anticipated that it would "be possible to make statistical analyses in far greater depth than could be made using only information that can be collected by direct retrospective enquiry" (OPCS, 1973 p.10). Although it was known that the range of variables from such a dataset would be narrower than that from other existing data sources (i.e. the General Household Survey), the accuracy of the data and its more regular collection and consistency through time were anticipated to be the main advantages. In addition to this, the sample size of the LS is also substantially larger than that from the General Household Survey (GHS). Therefore, an important part of the rationale behind the development of the LS was that the dataset would enable improved analysis of occupational mortality and fertility. Although, from its inception in 1971, the GHS included questions on birth intervals, sample sizes were too small for the level of analysis desired and the survey was retrospective in nature, meaning that information on past trends was being derived and required accurate recall and recording. The collection of further socio-economic information at birth through additional questions at registration would have required legislation.

Unlike many other datasets, the LS is not dependent on retrospective reporting of information, apart from at the census. Information on 'events' is collected on an ongoing basis and linked to the appropriate LS member through co-operation between the ONS and the NHS. It is this vital dimension to the dataset that makes this a longitudinal 'study' - those included in the dataset are unaware that they are included and have not opted into any survey or research. Being a 'study' means that non-response is not such a problem that it is in surveys, but creates complex problems in the tracing of individuals on the NHSCR and the linkage of individuals between censuses. These aspects are of key concern in this document and are explained further, as there are systematic trends in the linkage and tracing of individuals in the LS. Overall, the LS is very different to other longitudinal datasets in England and Wales because of its recording of events to individuals rather than reporting of events retrospectively. In terms of sample size, the LS is large compared to the British Household Panel Survey (BHPS) and has a unique form of adding and removing individuals. The size of the dataset and continued tracing of individual 'events' means that the LS is more comparable in its functioning with population registers elsewhere in Europe (e.g. Finland or Norway). Indeed, this makes the LS highly suited to comparative European research in areas such as fertility. Examples of work which has used the LS in such a way include that by Rendall (1999) and Rendall et al. (2005; 2009). The *Échantillon Demographique Permanent (EDP)* in France is similar in construction to the LS and has a comparable sampling fraction.

Members of the LS who die or embark (i.e. emigrate) have these events recorded or 'flagged' on their files. These variables can be included in the selection of individuals for analysis; members who have 'embarked' (left England and Wales by informing their NHS GP) or died are not removed from the LS, but remain in the dataset for analysis. In total there are over one million individual records in the LS, representing people who, at one time or another, have been LS members; but not all of whom have been traced or would necessarily be included in analysis. This is around twice the number of LS members that were sampled at the 2001 census and shows how many people have been followed since the inception of the LS. Emigrations or 'embarkations' are identified through de-registration with a GP. Given the size and long-term nature of the study, the accurate selection of a sample is a key step in the analysis of the data. LS members who have not been found on the NHSCR are 'untraced'. In some cases it is possible to link data for these members if they have been found in the LS 'no trace' index. These LS members untraced against the snapshot of the NHSCR from the census can be traced on the updated version of the NHSCR after the census (i.e. by the end of the census processing stages the LS member may have registered with a GP and be traceable on the NHSCR).

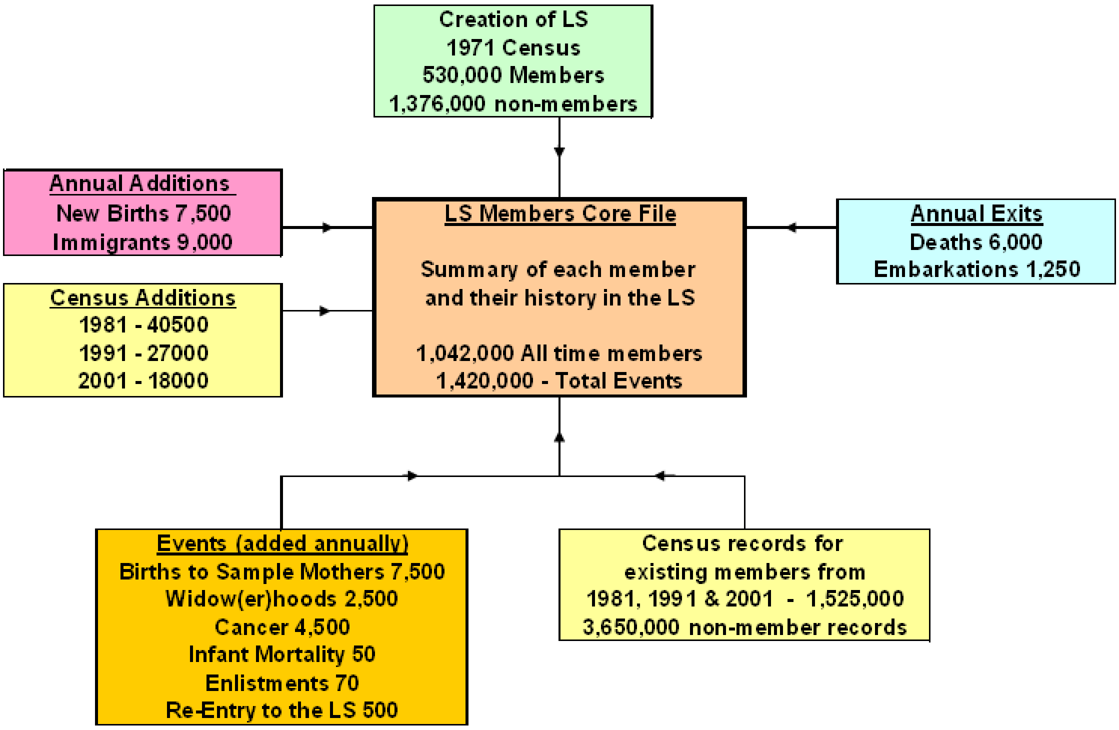
The potential identification of individuals in the LS is controlled through the Micro-data Analysis and User Support (MAUS) at the Office for National Statistics (ONS). Applications to use the LS are vetted and 'approved researcher' status required for research teams using the LS. To protect the identity of LS members, the inclusion of certain variables is restricted because of the ability to triangulate information and identify individuals (e.g. month of birth of an LS member and information about precise birth dates and cancer registrations might aid the identifications of an individual known to a researcher). Two forms of output can be approved by the ONS. The first of these are intermediate outputs where output with cell counts of more than 2 can be cleared from the Virtual Microdata Laboratory (VML), a secure computer setting at ONS offices. Final outputs are publishable materials, and cell counts of less than 10 are not permitted for publication.

3.2.2 Describing and understanding entry and exit points used in the Longitudinal Study

This sub-section explains the ways in which an LS member can enter the LS and ways in which embarkations occur. To say that a member leaves the LS would be incorrect as, although someone may die or embark, they are not removed from the dataset but no longer form part of the ongoing, traced part of the dataset which is linked to the NHS records. These LS members cease to be observed. Each entry and exit point from the LS is detailed and analysis made of the potential for problems with the use of that form of entry or exit. With the variety of entry and exit points, each has its own form of measurement and accuracy in capturing those that should be in the LS, given their residence status and date of birth.

Figure 3.1 is from the ONS LS Development Team and provides a schematic plan of the construction of the LS, along with broad numbers on the origin of LS membership. The initial sample (530,000 members), which was taken in 1971, is shown at the top of the diagram. Also shown are the 1.3 million 'non-member' cases that were taken from the 1971 census forms which included an LS member. These are normally 'co-residents' of the LS member and so individual-level information is potentially available for the people living with sample members. The 'co-residents' in the household of an LS member are not followed through time on the NHSCR or between census dates in the same way as LS members. All LS members (not co-residents) form the 'core file' which has annual additions (new births on an LS date and immigrations for those with an LS date of birth) and census additions. These are shown on the left of the diagram. To the right are the total number of annual exits (through death and embarkation) since the LS began. In total the LS core file has over a million 'all time members' (shown in the central box) which refers to those who currently are, or have been members, at one time or another. At the bottom of the diagram is the addition of annual events to the LS; note that, as discussed, a re-entry to the LS is considered to be an 'event' to an LS member. These events to LS members are different to the annual additions and annual exits which refers to the addition of LS members and the exit of LS members. The central 'LS Members Core File' is the basis for sample selection and analysis, as it links together all the sources of data and information on an LS member to give all the data available for each LS member.

Figure 3.1: Schematic plan of the construction of the ONS LS



Source: Office for National Statistics licensed under the Open Government Licence v.1.0. Reproduced from ONS LS Development Team; Webb, A., Harvey, J. and Hiscock, S., Presented at ONS LS Introductory Session 10-11-2009.

Entry to the LS

There are three ways of entering the LS:

- entering at birth by being born on one of the four LS birth dates
- migrating to England and Wales with an LS birth date and registering with a GP
- being present at a census with an LS birth date and not previously in the LS.

(Hattersley and Creeser, 1995).

In terms of timing of entry there are two main types - the first two on the above list are what could be identified as 'annual entries'; they occur regularly and are part of the ongoing, updated nature of the LS with the person becoming an LS member. The final form of entry on the list is one which only picks up those with an LS birth date every ten years, at the census. There is the potential for those entering at the decennial census not to have registered with a GP or, at the census to have given an erroneous date of birth and thus triggered entry to the LS. However, the inconsistency of the birth date with information on the NHSCR or the lack of a record on the NHSCR would mean that the LS member is not part of the 'traced' population in the LS. These persons are flagged in the LS. Non-traced records are of limited interest for research as no information from the NHSCR will be added to the LS files of the non-traced population.

Entry at birth

Births are entered into the LS automatically. All births in England and Wales must be registered by law and a 'draft entry form' is sent to the ONS Vital Statistics Output Branch (VSOB) by the registrar. The NHSCR is then notified of all births automatically by the ONS and updated, giving the child an NHS number. Those with an LS date of birth are identified as having such at the NHSCR processing stage and their record on the NHSCR is traced. There is very little scope for error here and the way in which the events are recorded and pass through the ONS mean that the potential for missing those persons who should be LS members is minimal.

Entry through migration

The addition of immigrants with an LS birth date to the LS is more complex. New immigrants are issued with NHS numbers and added to the NHSCR when they register with an NHS doctor. An immigrant is defined as someone who has described themselves as such and given a previous address which is not in England and Wales. "The category of immigrant includes not only those individuals who describe themselves to their general practitioners as such, but also those who, having quoted previous address abroad, cannot be matched to an existing NHS number" (Hattersley and Creeser, 1995 p.25). (These persons are sometimes referred to as 'New Flag 4s' on the Patient Register in ONS literature). Immigrants are those persons arriving from outside England and Wales, including Scotland, Northern Ireland and the Channel Islands. This is a relatively open condition; there seems to be no consideration of the length of residence or the future residence in England and Wales.

The inclusion of new migrants to England and Wales in the LS is achieved through the details of their registration with a GP. Registration with a GP is not compulsory and some migrants may only register with a GP if, and when, they require medical attention. There can be a considerable time lag between the actual migration to England and Wales and an eventual GP registration. Compared with the rolling register system in other continental European countries, this is a weaker element of the LS. A population register is advantageous in this situation as it would pick up an entry to the country because there is a requirement that the person registers with the authorities when they arrive (or move within the country). The LS depends on the migrant registering with an NHS GP shortly after arrival in England and Wales. Those persons that have entered England and Wales in the intercensal period and were born on an LS date, but have not registered with a GP, should be captured on a census form. However, with no health record created because of their lack of NHS registration, it is not possible to trace these LS members as there is no entry on the NHSCR. There is also scope for a time lag between entry to England and Wales and registration with a GP and entry on to the NHSCR. The entry of migrants into the LS is widely recognised as a problematic area: "The capture of immigrants in the LS using information generated at NHSCR is known to be unreliable" (Hattersley and Creeser, 1995 p.116). Although there is the potential for a lag between entry to England and Wales and registration with a GP (and thus entry into the LS), the LS seems to be better at recording moves into England and Wales than exits from the dataset.

Entry at a census

At each census there is a section asking for the birth dates for all those that are at that address on census night. This includes visitors to addresses (not at 2001) so as to capture those that are normally resident at another address but are present elsewhere on census night (termed 'multiple enumerations' as the individuals were 'enumerated' at more than one address on the census night). In turn, as part of the census processing it is possible to identify persons who were a 'multiple enumeration', and these are flagged on the LS. All those with an LS birth date at the census are extracted from the census file and links to those at the past census are made.

The census is the opportunity to add in those that have not been included through the other forms of entry as described above, particularly movement to England and

Wales through migration. However, if an eligible LS member is not included in the census then they would not enter the LS. More problematic is that if an eligible migrant has not registered with an NHS GP, then his or her record from the census form cannot be traced to an NHS record. This means that, although the person might become part of the LS, they are not part of the traced population for whom events such as births and cancer registrations can be linked. In the past there have been retrospective exercises to add in the non-traced population when an NHS record has been found.

Exit from the LS

As already outlined, although the term 'exit' or 'embarkation' is used to describe when someone leaves the LS, the full details which have ever been held in the LS for that member are retained and the exit recorded as an event. The record remains in the dataset and all data for an individual can be used for analysis taking in a particular time frame. This is one of the merits of a longitudinal dataset, particularly a study, of this type.

Exit points for LS members are:

- deaths - recorded by the NHSCR
- 'embarkation' - emigration from England and Wales recorded by NHSCR through de-registration with a GP (and data from other sources)
- entry to the army
- long-stay psychiatric hospital visits.

(Hattersley and Creeser, 1995).

Death

Deaths must be registered by law and data on deaths in the LS is said to be highly accurate. "The quality of death data for England and Wales is, like births data, extremely high. Death certificates are required by law before burial or cremation of a body, and as a result, virtually all deaths occurring in England and Wales are registered" (Hattersley and Creeser, 1995 p.117). Delays in certification can occur if a death occurs overseas, or if there is an inquest. The process used in the inclusion of deaths into the LS is identification of LS members by an annual computer file search for all deaths occurring to persons with an LS birth date and by flagging at the NHSCR (all deaths are routinely notified to the NHSCR). Where the deceased was born on an LS date, but there is no NHS record, then that record is removed from the LS files. Through notification to the NHSCR and the ONS there are two opportunities to identify LS member deaths. It seems that the dual system for the inclusion of LS members deaths is robust and that there is little opportunity for deaths to be missed.

'Embarkation'

More problematic is the accurate measurement of the embarkation of an LS member. The use of GP de-registration and the other sources of data (listed below) do not ensure that all migrations are captured. The numbers of people leaving England and Wales without de-registering, or having details recorded at the above listed organisations, must be substantial. It is only known who has embarked and not de-registered at the census when the person does not appear on a census form and cannot be matched. However, an LS member could be resident, yet was not recorded on a census form and therefore has not actually embarked.

Information on the assumed day of embarkation is received at the NHSCR from the following sources in addition to information on de-registration at a GP:

- Family Health Service Authority (formerly Family Practitioner Committees (FPC's)
- Ports Authority
- consulates
- embassies
- Department for Social Security (DSS).

(Hattersley and Creeser, 1995).

As a way of quickly understanding the problem of embarkations in the LS, it is possible to compare the number of people registered with a GP in England relative to the number of persons estimated as being resident according to ONS mid-year estimates. Simple background research on the numbers of people registered with an NHS GP in England shows a much higher number of people registered with a GP relative to ONS mid-year population estimates. Data for 2008 shows that there were 5% more people registered at a GP in England as a whole than resident in the population. To correct for this, the NHS uses a reconciliation technique which revises the NHS data towards the preceding year's ONS mid-year estimates (in this case data from 2007 ONS mid-year estimates was used). This is done in part because of the way in which funding is given to GPs and the need for accurate numbers. The geographies used for both the NHS data on GP registrations and ONS data on mid-year estimates are Strategic Health Authorities (SHAs). It is possible to use data at a more detailed geographic level, but the SHA level is sufficient for an overview.

Table 3.1 shows in rank order the percentage overcount of 'unreconciled' GP registrations compared with ONS mid-year estimates. London has the largest overcount, with 11% more people registered with an NHS GP than estimated resident in 2008 by the ONS. This is more than double the percentage overcount for England. Other areas where the number of people registered is, in percentage terms, far higher than the national average are the West Midlands and North West SHAs. Data on 'unreconciled' GP registrations is not available by sex. Given that it is not possible to be registered with more than one GP at any one time, the data suggests that there are areas where there are far more people registered with a GP than can actually be living there. It is unsurprising that London is the area with the highest over-representation of people registered; of all areas in England, London is the most 'globalized' in terms of immigrant communities, workers originating from overseas and international students. Not all people who have migrated to, and lived in, England for a period will have registered with a GP. Among those who have, it is likely that there are high proportions who have left the country without de-registering with their GP or had their 'embarkation' recorded by one of the other sources that the NHSCR uses. The data shows the likely excess numbers of people on the NHSCR register that are not actually resident. Note that in the North West NHS SHA there is still a higher number of persons thought to be resident after 'reconciling' the GP registrations to the mid-year estimates than compared with ONS mid-year estimates.

Table 3.1: GP registrations (2008) and ONS mid-year estimates (2008) for all patients, England

Strategic Health Authority Area	'Unreconciled' GP registrations	'Reconciled' GP registrations	ONS mid-year population estimates	Percentage over count of 'unreconciled' registrations compared with ONS mid-year estimates (2007)
London	8,463,608	7,562,531	7,619,800	11.07
West Midlands	5,723,506	5,374,319	5,411,100	5.77
North West	7,262,680	6,893,535	6,875,700	5.63
England	*53,944,734*	*50,868,539*	*51,446,200*	*4.86*
South Central	4,257,687	4,000,246	4,062,300	4.81
South East Coast	4,496,005	4,246,830	4,317,800	4.13
South West	5,378,671	5,128,860	5,209,200	3.25
North East	2,653,954	2,559,475	2,575,500	3.05
Yorkshire and the Humber	5,362,944	5,166,618	5,213,200	2.87
East of England	5,836,038	5,608,524	5,728,700	1.87
East Midlands	4,509,641	4,327,602	4,433,000	1.73

Own elaboration based on data from NHS Information Centre, Leeds, October 2009.

Re-entry to the LS

In addition to the forms of entry to the LS as outlined above, it is possible for a previous LS member to re-enter the study. Re-entries to the LS are from members who had 'left' the NHSCR by embarking, joining the armed services or being committed to a long-stay psychiatric hospital. When an LS candidate re-enters the LS this is recorded on their file and they are part of the 'traced' population in the LS again (NHSCR information is attached to them). The identification of re-entrants to the LS is relatively simple, in that these cases have an NHS number. Immigrant and re-entrant files come from the Family Health Service Authority and files for immigrants and re-entrants created at the NHSCR, which updates Central Health Register Inquiry System (CHRIS) by flagging LS immigrants and re-entrants (Hattersley and Creeser, 1995).

Although an appropriate system is in place for the capture of re-entries to the LS, there is the potential for LS members to return to England and Wales and not re-

register with a GP. The lag in return to England and Wales and re-registration could be considerable. Identification of a re-entered LS member at the census is possible, but there would not be an active NHSCR record which can be used for the tracing of an individual.

3.2.3 Events to LS members

The complete list of events which are recorded for each LS member is extensive. In full, events data is collected on:

- new births on LS dates
- births to LS sample members
- infant mortality of LS members' children
- deaths to LS members
- immigrants and re-entrants to the NHS
- cancer registrations
- death of the spouse of an LS member
- enlistments into the armed forces, embarkations, entries into long-stay psychiatric hospitals and re-entrants back into the NHS.

(Hattersley and Creeser, 1995).

The primary concern of this research is the accurate inclusion of births to sample mothers, as the addition of new LS members born on an LS date of birth has already been discussed above. A single system of identifying births is used, unlike the dual system for cancer registrations and deaths. From the birth registration process the parent's date of birth is used to identify LS members. These cases are then included in the annual births computer file and the draft entry form sent to the NHSCR where the LS is searched and the LS member's LS number is added to the draft. This is then sent back to the ONS and the LS number is used to link all data on the birth draft to the LS member. There is only one source of data which can be used, unlike the case of births and cancer registrations.

The miss-quoting of a date of birth is problematic, as this is the only way of extracting all those births that occur in a year. Although the accuracy of the linking process could perhaps be enhanced by using more details from each of the birth registration documents, the extra clerical review and complexity of the process involved might be considerable. A birth must be registered within 42 days of the birth in England and Wales. If the baby was born in England and Wales it must be registered in England and Wales (Office for National Statistics, no date c). The registration requirements of a birth are stringent and there is the opportunity for citizens of England and Wales to register a birth occurring overseas in England and Wales through the Foreign and Commonwealth Office (FCO) at the nearest consulate, or with the FCO consular department in London. The FCO will not register a birth if the parents were born overseas and are British only by descent. Non-British children adopted by British parents also cannot be registered (Foreign and Commonwealth Office, no date).

3.2.4 Initial conclusions on entry, exit and re-entry routes in the LS

Among datasets covering England and Wales, the LS is unique and offers detailed information similar to that in countries with population registers where every individual in the population is followed through time, and event information recorded. The linkage of the NHS record data to the one per cent sample in England and Wales provides an exceptionally large sample compared with other longitudinal studies (e.g. BHPS). As explained in this section, entry and exit points for LS members do not all function in the same way. Not all entry and exit events are recorded with the same degree of accuracy. Subsequent sections will deal with the relative accuracy of the LS for fertility analysis.

Although it is correctly claimed that there are fewer opportunities for attrition in the LS because it is a ‘study’ rather than a ‘survey’, the LS still has attrition which is often inadequately understood or explained in research using the data. This section began by outlining the rationale for the development of the LS in order to understand its intended purpose. Many of the founding reasons for establishing the LS are still important for researchers today. The LS is a large-scale dataset with a sampling frame which should lead to a broadly accurate representation of the population of England and Wales. With a survey, although there is more scope to leave, particularly in a repeated, longitudinal study like the British Household Panel Survey (BHPS), the way in which it operates is much simpler. The LS, through being a non-consensual study, means that there is no option to enter or leave, yet members are still lost through various mechanisms and, given the size of the dataset, these numbers can be sizeable. An individual entered into the LS at a census or through registration at a GP might not be identifiable at the next census for various reasons. Although no record of a death or embarkation might be recorded by the NHSCR and put on the LS members file, at some point in the intercensal period an unrecorded embarkation may occur. There are likely to be large differences with respect to the relative attrition of different groups in the population - this will be outlined in terms of 'tracing' and 'linkage' in the next section. In sum, although non-response might not be an issue, there are likely to be different people present at one census relative to the next.

Understanding of the various entry and exit routes which are used in the LS is the first step in beginning to account for when, where and who enters and exits the LS. The next section explains the way in which data is merged and linked to compose the LS as used by researchers. Later in this document hypotheses about the sources and likely impacts of attrition, through non-response at the census and non-recorded embarkation, are given.

3.3 Data used in the LS and an explanation of joining processes

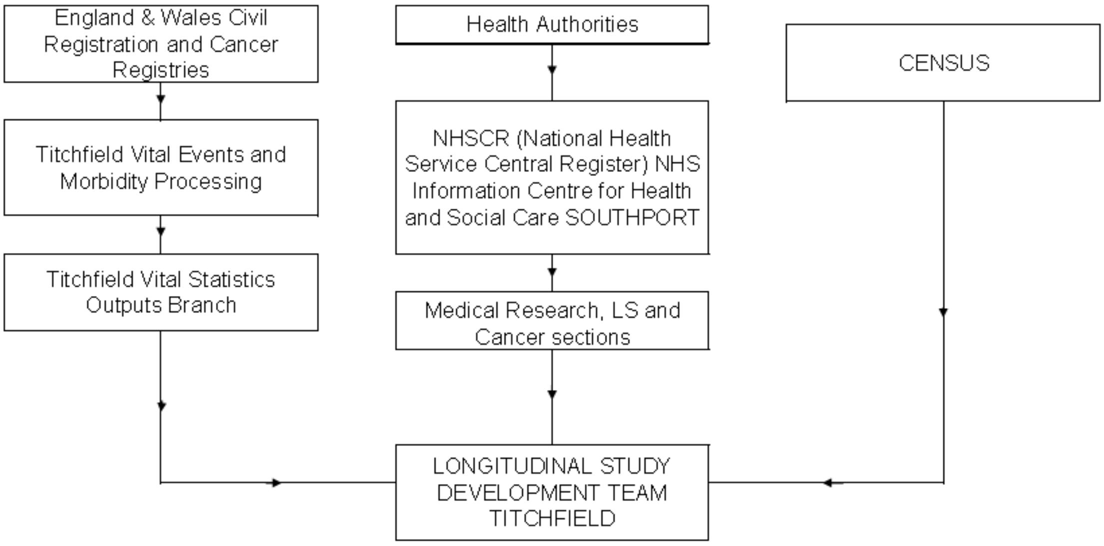
This section focuses on the different data sources used in the linking, tracing and sampling processes to create the LS. As was outlined in one of the original reports on the need and viability of a longitudinal study (OPCS, 1973), the data used to create the LS was all held within the Office for National Statistics, and not processed by any external organisations (this has subsequently changed with the NHSCR patient register now with the NHS and not within the ONS). Each of the streams of data used to produce the LS is collected separately for administrative reasons and not primarily for the LS. The three sources of data used are from the ONS decennial census, ONS vital statistics outputs and the National Health Service Central Register (which is part of the ONS). In this section each of the data sources is outlined, and this is followed by a discussion of the processes used to integrate the data into one dataset. Through understanding the relative accuracy of processing at each stage and the exact processes leading to the creation of the dataset, sources of error can be understood. This section first gives precise detail on the data which are used to compose the LS and then explains how these are linked to form the LS. In the latter part of the section, 'matching' and 'sampling' are introduced. These are important concepts for understanding the accurate combination of data and the continued following of individuals over time.

3.3.1 Data used in the Longitudinal Study

Figures 3.2 and 3.3 below have been provided by the LS Development Team, who are responsible for the construction of the LS. There are two ways in which 'event' information is added to the LS: first through a date of birth search on the annual computer files for England and Wales, and second through routine notification from the NHSCR. Figure 3.2 very clearly shows the three main strands of data of which the LS is composed. Taking each of the elements in turn is beneficial here. In the first column is data sourced from the ONS vital statistics output branch at the ONS. The Titchfield Vital Statistics Output Branch source of data in the first column is key for the addition of new members to the LS through births, the addition of birth information to LS members records and information on deaths. The NHSCR strand is shown in the second column, data which is used is that on new cancer registrations, embarkations, re-entries and other health information, as fully explained in the previous section. Importantly, the NHSCR is used as the 'hub' through which the accurate linking of information on LS members is completed. Census data is the third source for the LS and shown as the third strand on the diagram. This data comes from the decennial census and is linked to LS records using the NHSCR.

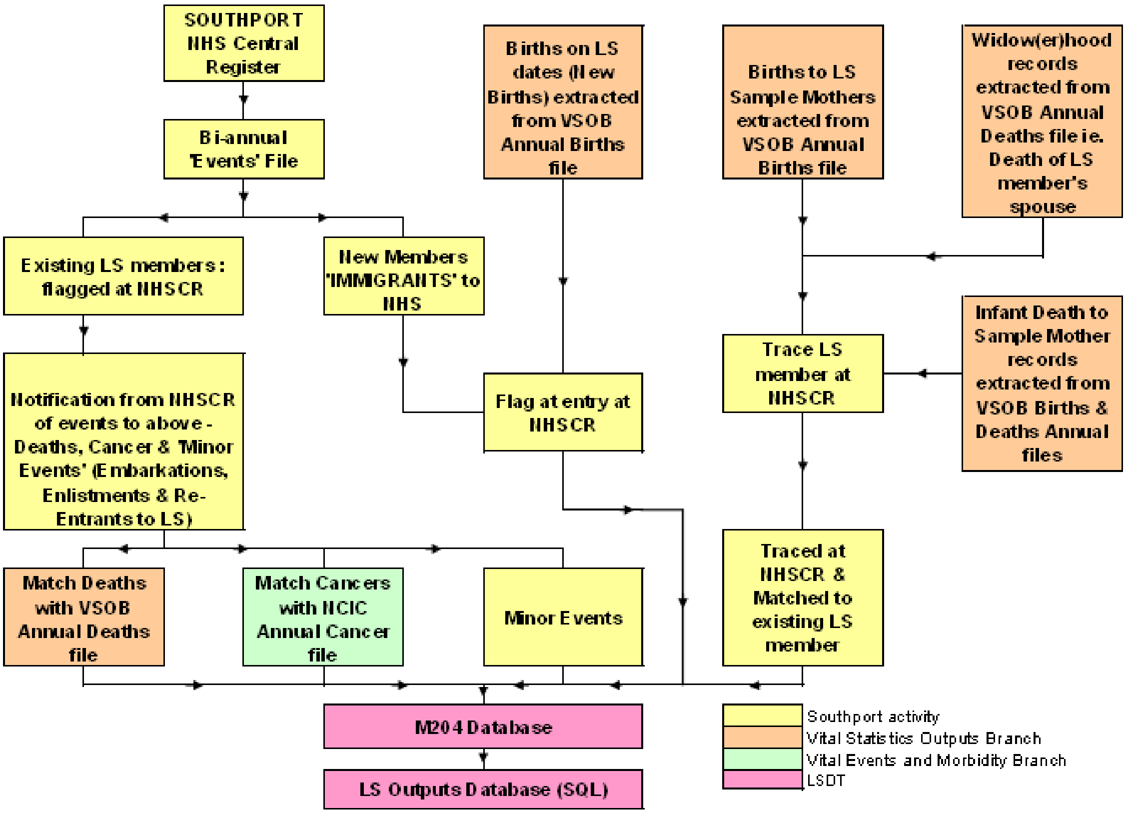
Figure 3.3 shows the level of complexity involved in putting together the different types and sources of data that are used in the construction of the LS. In this diagram the final outputs are at the bottom in the form of the 'M204 Database' and the 'LS Outputs Database (Structured Query Language (SQL))'. These two outputs draw all the above data sources together and provide a dataset with information on each LS member. In the figure, the degree to which much of the information comes from, or is at least processed at NHSCR in Southport, is clear. All of the data in the diagram has at least been processed there, including 'events' from the ONS vital statistics branch from the top right of the diagram.

Figure 3.2: Data sources for ONS LS composition



Source: Office for National Statistics licensed under the Open Government Licence v.1.0. Reproduced from ONS LS Team; Webb, A., Harvey, J. and Hiscock, S., Presented at ONS LS Introductory Session 10-11-2009.

Figure 3.3: Annual processing of events to compose ONS LS



Source: Office for National Statistics licensed under the Open Government Licence v.1.0. Reproduced from ONS LS Team; Webb, A., Harvey, J. and Hiscock, S., Presented at ONS LS Introductory Session 10-11-2009.

Census data - Office for National Statistics

The initial starting point for the LS was the 1971 census. In subsequent censuses responses for LS members have been added. Following each census, LS candidates have been identified where these persons have the same date of birth as one of the four LS dates. Through the use of the NHSCR, cases with inconsistencies between the two data sources can be linked and resolved. Figure 3.2 shows the way in which the census data is one of the strands of data used in the LS. However, data from the census is only available every ten years, which becomes more problematic the further one moves away from the last census, because the census is the only way of obtaining socio-economic variables for members of the LS. Changing socio-economic positions cannot be followed without the use of variables from the censuses. We only know the characteristics of the LS member at the census, cross-sectionally.

The most recent census for which data is available is the 2001 census, and at the current time linkage of 2011 census data is underway. In the post-2001 census processing stage an extract was taken from the 2001 census, containing the information on people who were eligible to be LS members required for tracing and linkage processes. At the 2001 census the use of electronic images of 'LS candidates' was used in the matching process for the first time. These were deleted after the processing stage to protect the identity of LS members. The extract was linked to the LS by tracing on the NHSCR and matching the individuals to existing LS records. An electronic matching process was used linking the census data to NHS data based on the name, date of birth and postcode of enumeration to identify exact, single matches which were accepted. Any other cases, including those where there was more than one candidate matching, were left for clerical review and matching. Those records not found on the CHRIS were checked manually (clerical review). The automatic matching at 2001 led to 74% of records being matched and, with clerical review following up the cases that could not be matched in the automatic stage, the final match rate was 96% (ONS, no date a). In addition to looking at the 2001 records, ONS staff also checked the 'no trace' file at NHSCR and took unmatched 1981 and 1991 records and re-checked them. Additional trace routes included searching the computerised 2001 Electoral Roll. The final census extract contained variables that had been subject to editing and imputation processes by the census division at the ONS, but did not contain the imputation used in the 'one number census' to create individuals thought to be missing at the census.

Vital Statistics data - Office for National Statistics

As is outlined in the first column of Figure 3.2, vital statistics outputs from the ONS Vital Statistics Output Branch are the second strand used in the creation of the LS. Events detected by the use of birth stated on an event document and picked up through a search of the annual computer files for England and Wales are:

- live births occurring on LS dates
- live and still births to women born on an LS date
- widow(er)hoods to LS members
- infant mortality (of births to LS members)

- deaths to sample members
- cancer registrations

(Hattersley and Creeser, 1995).

Event and health data - National Health Service Central Register (part of the Office for National Statistics)

The NHSCR is central to the successful operation of the LS. With no personal identification numbers or population register, the NHS number is the only commonly held means of identification. Every person who has registered with an NHS GP has an NHS number. In the post-census linking stages, one of the first steps is to link the census data to the NHSCR, which has all the necessary information for linkage to be made between the variables in the census and the NHSCR. The serial numbers for LS members can then be used to attach the new census information to LS members. In terms of the addition of event information, the NHSCR is also central to the successful operation of the LS, as the events in which the LS has an interest are routinely notified to the NHSCR. Routine notifications to the LS from the NHSCR include:

- enlistments into the armed forces
- embarkation or emigration (normally referred to as embarkations)
- reinstatements into the NHS (normally referred to as re-entries)
- new entrants to the LS: immigrants and re-entrants to the NHS, not previously LS members, picked up on stated date of birth
- deaths to sample members
- cancer registrations

(Hattersley and Creeser, 1995).

An important point to note is that deaths to sample members and cancer registrations are recorded through the use of routine notification procedures and linkage through the stated date of birth on event documents. This means that there is effectively a ‘dual system’ for the recording of these events in the LS, which minimises any chance that an event might be missed. Unfortunately, there is no such system for births - they are recorded just once through the Vital Statistics Output Branch annual files and then extracted and provided to the LS for linkage through the use of the NHSCR.

3.3.2 The processing and joining of data to produce the LS

The need to join the three sources of data has already been discussed in the preceding text. This section outlines the way in which the data sources are joined to create the LS, and identifies where there is the potential for data quality problems. The key concepts and terminology are:

- *Matching* - refers to the matching of a census record for a person born on an LS date of birth to a corresponding record for that person in the LS. This is done at the NHSCR. If a record is matched, then that person can be studied longitudinally. The measurement of the matching of LS members' records from one census to the next is done through linkage rates, which refer to the proportion of LS members found at a census who were resident at the preceding census. Individuals will not be identified if they were not at a census, information on their death or embarkation was not recorded on the NHSCR, or there are inconsistencies between dates of birth or other personal information given at censuses and used for linkage.
- *Tracing* - attachment of NHSCR events to LS members. This is done through identification of LS members on the Central Health Register Inquiry Service (CHRIS) in Southport. Data from the census is provided by the LS team at Titchfield. The NHSCR enables records for LS sample members to be linked to various life events for these individuals. In quality terms, 'tracing rates' are discussed, and indicate the likelihood of both census and event data linkage for groups within the LS. LS members are 'untraced' if they have not been found, either because they have not been registered with a doctor, or inconsistent names or dates of birth have been used.

The next sub-sections explain the terms above and review reports on each of these aspects which are important to data quality. As outlined in the last section, there is the potential for data quality problems with regard to the 'tracing' of LS members and the attachment of event data from the NHSCR to them (births being the concern here) and the 'linkage' of individuals between the census dates. Periodically, the ONS produces reports and statistics on these concerns and these are analysed under each of the above terms.

Another term which must be considered is the 'sampling' of the LS in relation to the population of England and Wales. As was explained, the target sampling fraction of the LS is 1.09% - this is the proportion of the total population of England and Wales that should be in the LS. The following term explains the sampling of the LS. The representativeness of the LS in relation to the population of England and Wales is best understood in terms of the sampling fraction.

- *Sampling fraction* - the degree to which the LS represents the actual population as calculated from vital statistics. Sampling fractions are calculated using traced LS sample members and census populations without adjustment for census under-enumeration. The traced LS sample is divided by the census population and multiplied by 100 to give a percentage.

The sampling of the LS is discussed below, in a section which includes analysis of reports by the ONS on the sampling of the LS.

'Matching' of LS members and linkage rates

'Matching' and 'tracing' are processes in the construction of the LS which could be easily confused. It is crucial to remember that matching is concerned with finding an LS record for a person born on one of the LS birth dates, while tracing is concerned with finding a corresponding NHSCR record. Matching begins with taking a new census record for a person with an LS birth date and finding a corresponding record in the LS. Matching of LS members in the post-2001 census processing stages was made on the following characteristics, according to the census form:

- name
- date of birth
- sex
- postcode

- postcode one year ago
- person type (private / communal establishment)

(Office for National Statistics, no date d).

Linkage rates refer to the proportion of members of a given group (e.g. age in census year) who are successfully matched at a census to an LS record at the last census. 'Valid exits' in the form of deaths or recorded embarkations (via de-registration at an NHS GP) are taken into account in the calculation of linkage rates. Individuals are not identified (matched) if they were not at a census, information on their death or embarkation was not recorded on the NHSCR, or there are inconsistencies between dates of birth, or other personal information used for matching given at censuses.

A review of ONS publications on the census to census linkage of LS members on the NHSCR

After each census a report on the matching rates achieved by the LS is published by the ONS (the post 2001 report was by Blackwell et al., 2003). This details the linkage of individuals between the census dates and the characteristics of those where there were persistent linkage problems. At the 2001 census ONS reports state that 88% of individuals were followed between 1991 and 2001. Linkage rates are low among young males, those never married, those divorced or living in a lone-parent household, ethnic minorities, those living in a communal establishment, unemployed, or a student and those in the armed services. In many respects these are what could be called 'the usual suspects' in terms of problems of representation with census data. Another key indicator is the linkage by age group, and this shows that people in the 20-29 year age group and those aged over 75 years in 2001 were where failure was most likely. Economic position was also key in determining appearance at the 2001 census for LS members - linkage failure was most prevalent among those waiting to start a job, the unemployed and those on government schemes. Those born in the UK were more likely to be linked in 2001, compared with those born elsewhere.

In literature on the operation of the LS the term linkage is used to describe the process of following individuals between the census dates. This is perhaps the most crucial aspect of the operation of the LS as a whole - without the accurate linkage of individuals from one census to the next the longitudinal nature of the study is weak because too few individuals can be followed through time. At each census all those people found on a census form with an LS date of birth, and who are usually resident in England and Wales, are extracted from the census households file. This includes 'multiple enumerations', persons recorded at more than one address on the census night (e.g. people who were counted twice by being recorded at one address on the census night but being usually resident at another address, which was given on the census form). 'Secondary records' for the multiple enumerations are recorded in a separate file, yet there is a flag available on their main record in the file so that they can be identified.

The 2001 census processing stages included imputation of missing households, individuals and non-response for specific questions. However, for the LS there was a need to use non-imputed individual records (as these would not be traceable on the NHSCR) and identify where any imputed values were used for non-response in specific questions. The 2001 census used a form of imputation termed Edit and Donor Imputation System (EDIS), which was used to fill in all the gaps for existing people and households (where there was incomplete information as a base for a person). The use of non-imputed individuals is an important point - the LS would have higher non-trace rates if there was insufficient information to trace members on the NHSCR, or to match them over time. LS members were not used as 'donors' for the imputation of other values or records in the census data. Data from the census extract was supplied to the NHSCR, who traced those present at the census with an LS birth date in NHSCR records. Where possible, a match to existing records was made. Inevitably, it was not possible to trace members on the NHSCR or match to an existing LS member file in all cases.

Table 3.2 shows the percentage of LS members found in 1991 but who were not at 2001 by age and sex. The data shows that there are problems with the representation of women in the main childbearing years; there is no age group under the age of 40 years where there was more than 90 per cent linkage success. For men in the LS the figures are worse, with a quarter of those aged 25-29 years missing at the 2001 census, either through an unrecorded embarkation or not being recorded on a 2001 census form. Overall, 10.7% of women who were at the 1991 census are missing in 2001. In historical terms, the third decade was a bad one for linkage in the LS - the rates described here are lower than at 1991 or 1981. In sum, this means that using residence at the 1991 and 2001 census dates as a selection criterion reduces the number of members for analysis. A finer-grained approach to sample selection would be beneficial. Benchmarking to other population statistics on offer from the ONS would be helpful for further understanding the longitudinal nature of the LS given cross sectional variations at the census.

Table 3.2: Post 2001 census linkage summary

Traced LS members found in 1991 but not accounted for in 2001, by age and sex				
Age in 2001 Years	Males		Females	
	Found in 1991, not recorded as having died or embarked by 2001	Not found in 2001 (%)	Found in 1991, not recorded as having died or embarked by 2001	Not found in 2001 (%)
10-14	17,955	11.6	17,424	12.0
15-19	17,043	12.9	16,261	12.7
20-24	16,390	21.2	15,608	16.5
25-29	17,556	24.1	17,107	16.0
30-34	18,989	21.5	19,711	13.9
35-39	20,230	16.5	21,173	11.7
40-44	19,192	14.7	19,418	9.9
45-49	17,613	12.1	17,759	9.1
50-54	18,820	9.9	19,523	7.8
55-59	15,823	9.1	16,068	7.7
60-64	13,653	8.8	13,874	7.7
65-69	12,082	7.5	12,843	6.8
70-74	10,399	7.6	12,122	6.9
75-79	7,939	7.4	11,075	7.5
80-84	4,534	8.8	7,904	9.4
85-89	2,284	14.4	5,215	13.3
90-94	687	16.3	2,179	11.5
95-99	119	16.0	569	18.5
100+	13	38.5	83	26.5
All ages	231,321	13.8	245,916	10.7

Source: Office for National Statistics licensed under the Open Government Licence v.1.0. Reproduced from Blackwell et al. (2003) Longitudinal Study 1971-2001: Completeness of Census Linkage, Series LS no. 10.

'Tracing' of LS members

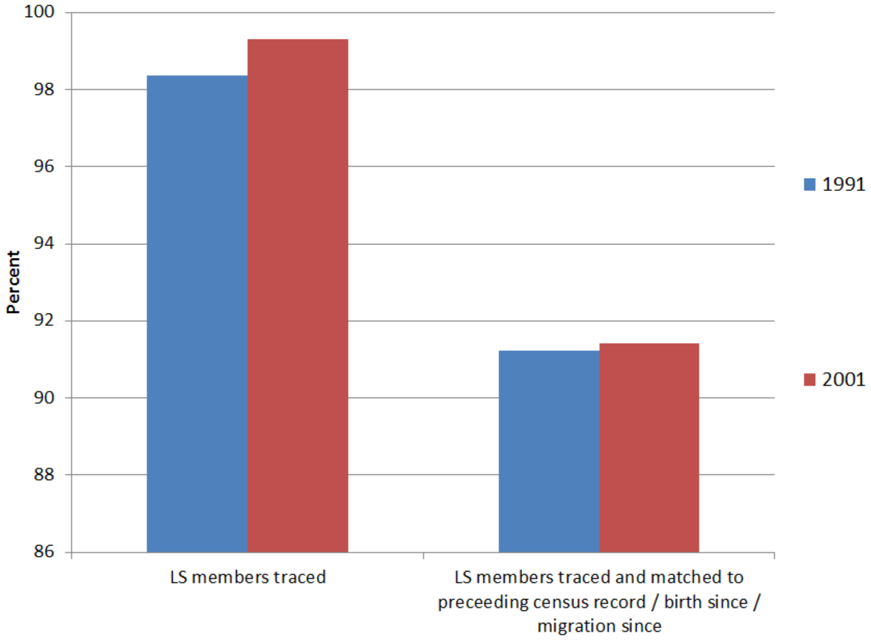
Tracing is an exercise that is carried out in the post-census processing stage and when a member enters the study for the first time. This is an important part of the construction of the LS, as it is only those traced LS members where the attachment of events data can be made. The NHSCR should contain a record for all members of the population of England and Wales who are registered with a GP, although immigrants may show a lag in registration with an NHS GP or not register, and there may be a lag in the removal of NHS records that are no longer active. Although some of the data used in the annual events to LS members is not from the NHSCR, the NHSCR is used to attach the event information to traced LS members. For example, all births to women with an LS date of birth are sent to the NHSCR for attachment to an LS members record. The NHSCR record is used in the attachment process and an LS member must have an NHSCR record for events to be linked. A variable on tracing is available in the LS to identify at which point in time an LS member has been traced in the LS.

A review of ONS publications on the 'tracing' of LS members on the NHSCR

Reports on the tracing of individuals are produced after each census and for the 2001 census (Office for National Statistics, no date a.) it was reported that 99.3% of LS members were traced - this being higher than previous census dates. Trace rates were higher for women than men, but there were lower rates for those in the 20-24 year age range with 2% not traced at the census. Those lower rates for women in their early twenties link to the characteristics of those in the report that were least likely to be traced - being a young adult, being single, being born outside the UK, being in the economic position 'other inactive' or a full-time student and living in certain communal establishments. At the socio-economic extremes of the National Statistics Socio-economic Classification the non-trace rates were higher; the economically inactive and those in the higher professional occupations had the lowest trace rates.

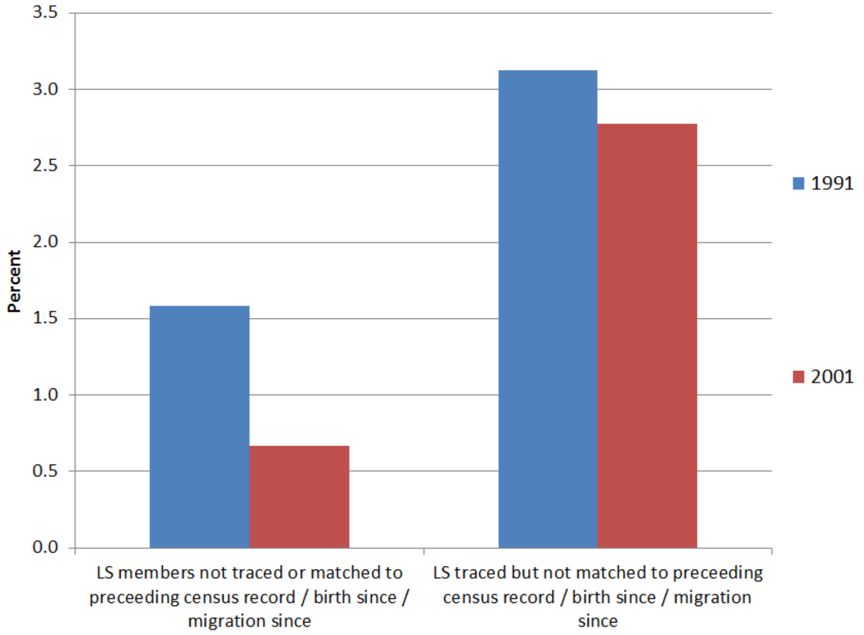
Figure 3.4 and Figure 3.5 show the matching and tracing histories for 1991 and 2001 respectively. The figures show that there has been an improvement in the percentages of LS members matched and traced to 1991 census records at the 2001 processing stage, compared with the 1991-1981 processing. There was little difference in the percentage of LS members who were traced at the NHSCR but who were not matched. Although the match and trace details included are of interest, there is no disaggregation by sex or age groups.

Figure 3.4: Successful tracing and matching of LS members at 1991 and 2001 to a LS record at the preceding census



Own elaboration based on 'Linking census data to the LS' (matching and tracing summary), Office for National Statistics, no date d.

Figure 3.5: Unsuccessful tracing and matching of LS members at 1991 and 2001 to a LS record at the preceding census



Own elaboration based on 'Linking census data to the LS' (matching and tracing summary), Office for National Statistics, no date d.

'Sampling quality' of LS members

Sampling fractions show the extent to which the LS represents the population of England and Wales. By taking the number of persons of a particular sub-group in the LS and dividing this by the corresponding number according to ONS mid-year estimates or vital statistics, it is possible to understand where the LS under- and over-represents certain groups. As outlined, the target for the LS is 1.09% of the population as a whole. However, this varies across groups. In addition to being concerned about the degree to which linkage is accurate in the LS, this work is interested in the sampling quality of the LS through time.

A review of ONS publications on 'sampling quality'

ONS reports (Office for National Statistics, no date b.) include data on where the LS over- and under-represents the population of England and Wales. Again, the sampling quality of the LS is discussed year on year for the LS as a whole. Table 3.3 shows the overall representation of new LS members entering via being born on an LS birth date since 1991. Of principal interest is the column labelled 'Sampling

Fraction' under the 'Total' column on the extreme right. This shows the year-to-year variations in the sampling of LS members born into the LS. The figures show that there is a fluctuation around the target fraction of 1.09. The sampling fraction is arrived at by dividing the number of births recorded in a year in the LS by the corresponding figure for England and Wales and multiplying this by 100. The totals provided at the bottom of the sampling fraction column should be used with some caution, as these are calculated using an average of the annual figures for each year. The table uses the term 'entry rate', which is normally called 'linkage rate'. To calculate the 'entry rate', the number of entries from being born on an LS birth date are divided by expected births in the LS. The number of births expected in the LS is calculated by dividing 365.25 (days of the year inclusive of 0.25 for leap years) by 4 (the number of birth dates used in the LS) and then multiplying this by the England & Wales births figure. In the third decade (1991-2001) there is a fairly even profile to the sampling fractions at just over 1%, until the middle of the decade where there is a rise to 1.18% and 1.15%. Following this, there is a decline to the latter part of the decade.

Table 3.3: New LS member entries through births, by sex 1991-2005

Third decade new births by sex and year of birth

Year of birth	Males					Females					Total				
	LS births	E&W births	Sampling Fraction	Exp. In LS	Entry rate	LS births	E&W births	Sampling Fraction	Exp. In LS	Entry rate	LS births	E&W births	Sampling Fraction	Exp. In LS	Entry rate
Part 1991	3,133	249,412	1.26	2,946	106.35	3049	237,166	1.29	2,801	108.85	6,182	486,578	1.27	5,747	107.57
1992	3,659	353,694	1.03	3,866	94.65	3475	335,962	1.03	3,672	94.64	7,134	689,656	1.03	7,537	94.65
1993	3,588	345,835	1.04	3,790	94.67	3353	327,632	1.02	3,590	93.40	6,941	673,467	1.03	7,380	94.05
1994	3,464	341,321	1.01	3,741	92.60	3439	323,405	1.06	3,544	97.04	6,903	664,726	1.04	7,285	94.76
1995	3,834	332,188	1.15	3,640	105.33	3792	315,950	1.20	3,462	109.53	7,626	648,138	1.18	7,103	107.36
1996	3,828	333,490	1.15	3,645	105.02	3653	315,995	1.16	3,453	105.79	7,481	649,485	1.15	7,098	105.40
1997	3,585	329,577	1.09	3,612	99.25	3369	313,518	1.07	3,436	98.05	6,954	643,095	1.08	7,048	98.67
1998	3,232	325,903	0.99	3,572	90.48	3103	309,998	1.00	3,397	91.35	6,335	635,901	1.00	6,969	90.90
1999	3,217	319,255	1.01	3,499	91.94	3079	302,617	1.02	3,316	92.85	6,296	621,872	1.01	6,815	92.38
2000	3,533	309,625	1.14	3,384	104.40	3443	294,816	1.17	3,222	106.86	6,976	604,441	1.15	6,606	105.60
Part 2001	878	99,319	0.88	835	105.20	841	94,548	0.89	795	105.85	1,719	193,867	0.89	1,629	105.52
Total	35,951	3,339,619	1.08	36,529	98.42	34,596	3,171,607	1.09	34,688	99.74	70,547	6,511,226	1.08	71,217	99.06

Fourth decade new births by sex and year of birth

Year of birth	Males					Females					Total				
	LS births	E&W births	Sampling Fraction	Exp. In LS	Entry rate	LS births	E&W births	Sampling Fraction	Exp. In LS	Entry rate	LS births	E&W births	Sampling Fraction	Exp. In LS	Entry rate
Part 2001	2,713	205,316	1.32	2,504	108.35	2528	195,451	1.29	2,384	106.06	5,241	400,767	1.31	4,887	107.23
2002	3,641	306,063	1.19	3,354	108.56	3576	290,059	1.23	3,179	112.49	7,217	596122	1.21	6,533	110.47
2003	3,442	318,428	1.08	3,490	98.62	3345	303,041	1.10	3,321	100.72	6,787	621469	1.09	6,811	99.65
2004	3,214	328,340	0.98	3,588	89.58	3106	311,381	1.00	3,403	91.27	6,320	639721	0.99	6,991	90.40
2005	3,543	330,600	1.07	3,623	97.79	3483	315,235	1.10	3,455	100.81	7,026	645835	1.09	7,078	99.27
Total	16,553	1,488,747	1.11	16,559	99.96	16,038	1,415,167	1.13	15,742	101.88	32,591	2,903,914	1.12	32,300	100.90

Source: Office for National Statistics licensed under the Open Government Licence v.1.0. Reproduced from New births into the Longitudinal Study, Office for National Statistics, no date e.

Table 3.3 shows that since 2001 there has been a good representation of new birth entries into the ONS LS, with only 2004 showing a lower rate of representation (under 1%).

3.3.3 Conclusions and key considerations for research design using the LS

The Longitudinal Study for England and Wales is a highly detailed dataset and spans a long time period, now incorporating four census dates and having a sample of over one million records (individuals that have ever been included in the LS since 1971). There are three main quality aspects for research using the LS to consider. As outlined above, these relate to the 'matching' of LS members between census dates, the 'tracing' of LS members on the NHSCR, and the 'sampling' of LS members which refers to the representativeness of the LS of the population as a whole. A key point to be noted here is the terminology which has become conventionally used in understanding the LS. 'Tracing' is used where 'linkage' is often used in demographic literature: to describe the process of attaching event information to an individual being followed over a continuous time period.

In general, the way in which persons are added to the LS seems reliable. However, the recording of migrations into and out of England and Wales is a problem area. "Recording of births and deaths is very reliable in the UK. However, notification of migration to and from England and Wales is not complete. This reduces the representativeness of LS sample estimates at time points between censuses" (ONS, no date f). This creates difficulties in particular for incomplete decades (i.e. between 2001 and 2011). However, using all the individual data which is collected on an individual to construct residence trajectories for their time in the LS would enable a clear understanding of the number of LS members who are exposed to risk of birth.

Literature from the ONS takes, as would be expected, a macro-level view - there are details of the groups that are under-represented and these are fairly consistent through time. Of major interest for the research area that has been outlined in Chapter 2 is the representation of births from LS members. The 'sampling quality' data could be expanded, with more information on births by age group and births to foreign-born women and comparability to other national statistics data given. This would provide a finer-grained perspective on the representation of LS births and allow a clear understanding of where births are over- and under-represented in the data. It is noticeable that, within recent published work using the LS, there has been no such detailed research.

Although there are reports on the linkage, tracing and sampling of the LS there is a need to further explore the trajectories of LS members in the third decade of the study. This is the last complete decade of the study, with the 1991 and 2001 census providing the start and end points. For accurate demographic analysis in the third decade a fine-grained and detailed approach to sample selection is needed to ensure the characteristics of the sample selected are fully appreciated. For analysis using the LS after the 2001 census the trends over the 1991-2001 period give an indication of those that are likely to be playing out in the period after 2001. Although, the higher levels of migration and ability of the LS to collect these new migrants is important to consider.

The detail of how the dataset is constructed and the background on the sample in past work using the LS is crucial for the next steps in assessing the use of the LS for fertility research. Indeed, the next section develops understandings on LS data quality through reviewing research on an individual basis to understand the ways in which data quality and sample selection has been approached.

3.4 What does past fertility research tell us about Longitudinal Study data quality and sample selection?

In addition to the ONS reports and background on the quality of the LS, it is beneficial to consider what previous fertility research using the LS has found on data quality and how sample selection has been approached. Although there is a wealth of research which has used the LS, approaches to data quality have varied, as will be explained in this section. To understand the approach of each piece of research to data quality, this section is structured around key pieces of individual research which have used the LS.

Babb and Hattersley, 1992 – *'An examination of Office of Population, Censuses and Surveys (OPCS) Longitudinal Study data for use in fertility analysis'*

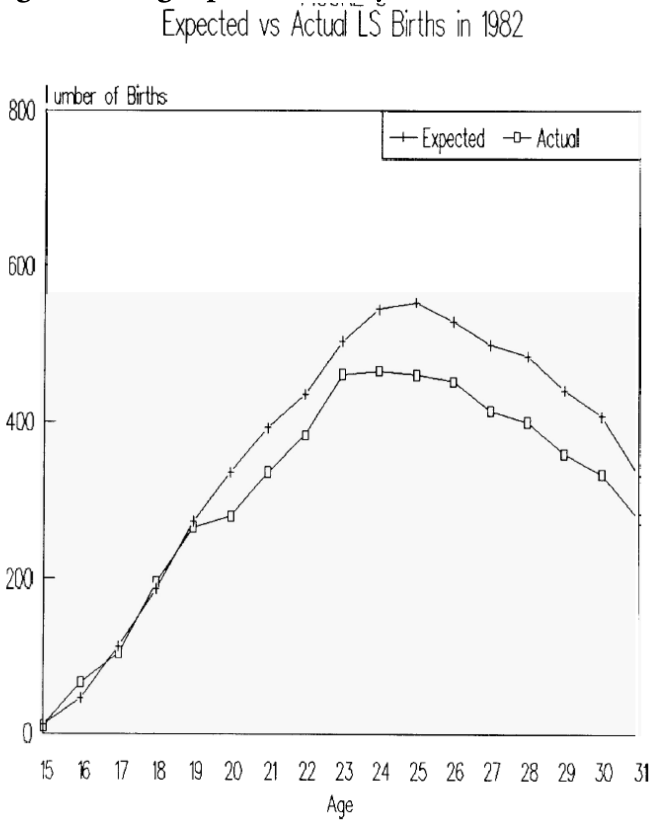
This research has aims which are broadly similar to those of this work. The work is titled 'An examination of the quality of OPCS Longitudinal Study for use in fertility analysis'. However, this is from the early 1990s and considers births in the period between 1971 and 1988 to women born after 1 January 1950.

Comparisons are made to ONS official statistics and figures from the General Household Survey (GHS). Throughout the document it is stated that the fertility rates are comparable and that the data is particularly appropriate for parity analysis because non-marital fertility is included. This is an advantage that is not as important as it once was, because other surveys no longer base their categories on only those that have had a birth in marriage. In the paper the improvents made to the LS in the 1980s are clearly explained and shown. Sampling fractions are calculated for women in the LS and also births to women in the LS, thus comparing the numerator and the denominator for fertility calculations. Babb and Hattersley found a greater variation in sampling fractions for births to LS members than they did for the actual numbers of women in the LS relative to official statistics. Sampling fractions had lower variation for births to older members in the sample and higher variation to the youngest women in the sample (15-17 year olds). Somewhat unexpectedly, given the reports that are detailed in the section on outputs from the ONS, older age groups are found to have lower sampling rates. The 15-17 year age group has a high rate – this is probably because of the small numbers involved. Linkage rates for births to LS members were compared to expected rates from the age-specific fertility rates (ASFRs) applied to numbers of LS women at each age. A correction factor was applied to the data to make up for the differences. With regard to period total fertility rates, it is found that there is a widening disparity from the 1981 census. It is highlighted late in the paper that the most fruitful analysis of the data is from a cohort perspective.

What are said to be acceptable variations in the data from the official figures seem in the graphs showing ASFRs to be quite large variations.

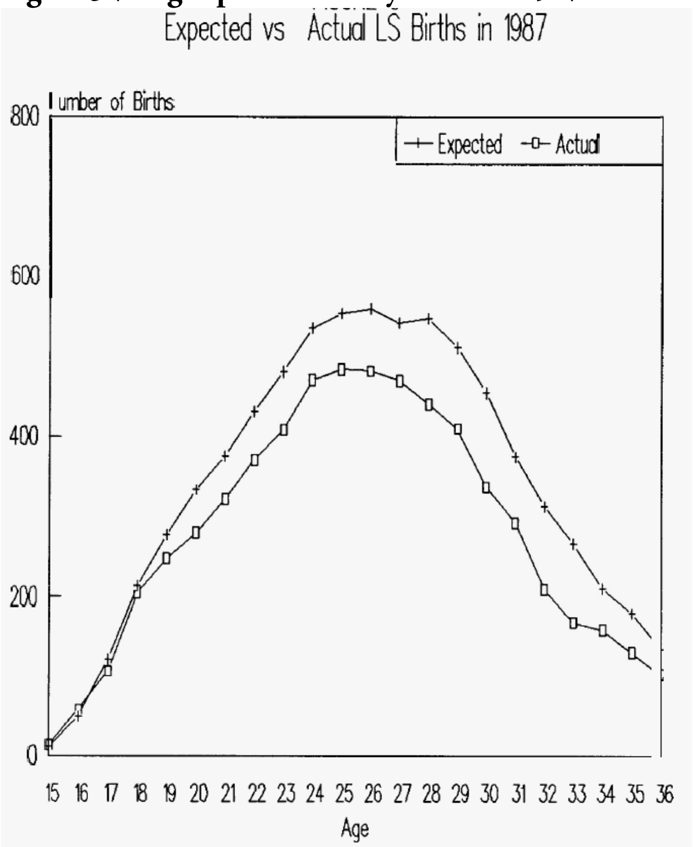
Figures 3.6 and 3.7 show the ASFRs which were calculated for the years 1982 and 1987 using the ONS LS in comparison with official statistics. Figure 3.6 (1982) shows that, although the actual sampling fractions for births to teenage LS members are lower than would be expected, there is a good ASFR which can be achieved using the LS compared to official statistics. At age 20 years there is the emergence of a gap which is fairly consistent into the mid-twenties, after which there is an expansion with an increase in the size of the gap between the official and calculated rates. Interestingly, Figure 3.7 (1987) shows that there is a lower level of comparability between the two sources of data than was the case in 1982. There is a slight change in the age at which the ASFRS are no longer directly comparable - at age 19 years there is the opening of a gap between the two data sources. Following this, there is a consistency in the gap between the age groups until the late twenties, where there is a slight increase, and the early thirties, where an increase can also be observed.

Figure 3.6: Age-specific fertility rates for 1982 from Babb and Hattersley, 1992



Source: Office for National Statistics licensed under the Open Government Licence v.1.0. Reproduced from Babb and Hattersley (1992), p.34.

Figure 3.7: Age-specific fertility rates for 1987 from Babb and Hattersley, 1992



Source: Office for National Statistics licensed under the Open Government Licence v.1.0. Reproduced from Babb and Hattersley (1992), p.35.

Hattersley and Creeser, 1995 – *'Longitudinal Study 1971-1991: History, organisation and quality of data'*

This report includes an extensive background on the reasons for starting the LS and points out that fertility research is one of the key reasons for its inception. The quality of linkage for LS members between the census dates in the report is said to be strong - 90% or more. A very helpful explanation of the two parts in event linkage is made – event recording and sampling, or linkage to the LS. Non-traced rates are said to vary substantially, with high rates for females aged 20-24 years. There are also large variations for those born outside of the UK and at the 1991 census. The highest non-traced rates were for persons born in the USA (potentially related to armed service personnel in England and Wales) and persons born in the West African Commonwealth and those born in countries given as 'the Rest of the World'. This report provides a wealth of detail on the first two decades of the LS (1971-1991) but since 1991 there have been changes in sampling rates and the linkage of new births to sample mothers.

Ekert-Jaffé et al., 2002 – *'Timing of Births and Socio-economic Status in France and Britain: Social Policies and Occupational Polarization'*

This more recent research aims to assess how fertility and the timing of births are linked to socio-economic factors in both countries, if 'social polarization' in Britain can explain the greater dispersion of births over the life cycle, and how the influence of socio-economic factors on fertility evolved on a cohort basis. The data sources used are the *Échantillon Demographique Permanent (EDP)* and the LS. Three cohorts are used in each country - those in France are from 1952-56, 1957-61, and 1962-66; England and Wales cohorts are from 1954-58, 1959-63, and 1964-68.

Within an annex to the main body of work there are some details on data quality. There is, however, more detail on the EDP than the LS. For the LS, the quality of links between census data and birth registration data is said to be acceptable, but the specific figures are not included. In the EDP around 10-12% of births are said to be missing, a figure which is similar to the rate in the LS, according to the work of Babb and Hattersley (1992). However, Babb and Hattersley (1992) are referring to the first two decades of the LS (1971-1991) and since then there has been an increase in the number of births linked to sample members. Among younger age groups it is identified that the omission rate is higher, especially among 'adolescents'. Correction for the missing data was made using other information on birth records (there is not precise detail on this). In total, the omission rate is said to have been reduced to between 5-6%, but with the EDP still under-estimating fertility. In outlining this, the authors argue that the groups which fall into this category are those that are in lower socio-economic groups who have their children at younger ages, so the bias resulting from the omission would not change the commentaries as the differences would be greater than those found. There is little detail on the methods used to correct for the problems identified.

Rendall and Smallwood, 2003 – *'Higher qualifications, first-birth timing, and further childbearing in England and Wales'*

The article is a study of the association between obtaining higher education qualifications and entry into childbearing. Higher parity childbearing is also studied. Women born in England and Wales between 1954 and 1958 are used in the study. The key finding is that the average age of entry to motherhood is five years later for women with higher qualifications than for those without. Higher parity births are fewer among those with a higher level of education.

The birth probabilities in the dataset were adjusted to national population statistics. Logistic regression outputs were 'corrected' for an "overall downward bias in their levels that arises through incomplete linkage of registered births in the dataset" (Rendall and Smallwood, 2003 p.20). In this case, by using the term linkage, Rendall and Smallwood (2003) are referring to the linkage of births from vital registration data to LS members. The correction changed the annual birth probabilities to a higher level to match the national population statistics rates on cohort, age and parity fertility.

"Ratios of these annual birth rates to the predicted annual birth probabilities from the regression (the latter summed over the two education groups according to their sample proportions by age and parity) were first calculated. These ratios were then applied to the regression-predicted annual birth probabilities by age, parity, education and duration since last birth." (Rendall and Smallwood, 2003 p.20).

Annual birth probabilities in the LS and GHS were compared by education and parity. Overall similar patterns of difference were found in the data between the two, and these were between women with and without higher qualifications and between women giving birth earlier and those giving birth later. So the results that were presented were interpreted as being unbiased with regard to making comparisons by education and age at previous birth. Longer birth intervals were found in the LS data, compared to the GHS data, and ratio corrections do not adjust for this, so the speed of progression to the next birth is said to be under-estimated uniformly

Overall, the corrective action taken is clearly detailed. It is recognised that there is a tendency for the LS to undercount births and the use of national statistics data to make a correction to the LS data ensures that the overall figures using the LS are higher. This transparency is reassuring and as a methodology for research using the LS and correcting for some difficulties with the data, this is a method which might have been suitable for other studies using the LS.

Rendall et al. 2005 – *'First births by age and education in Britain, France and Norway'*

This article in *Population Trends* looks at the progressively later starting of childbearing across Europe and makes a comparative study of Britain, France and Norway. Use of three, comparable longitudinal studies is therefore made. The focus is on age at entry to motherhood by education level among women born in the 1950s and 1960s. The specific cohorts used were women in born in the years 1954-58 and 1964-68 in Britain, women born in the years 1955-59 and 1963-67 in France and women born in the years 1955-59 and 1965-69 in Norway. On a comparative level, women born in Britain were found to have lower first birth rates in their mid-to-late twenties, relative to the late teens and early twenties. In France and Norway there was a stronger shift in first childbearing between the two age groups. Overall, in Norway and France the risk of first childbearing shifted more uniformly across education levels between the two cohorts used than in England and Wales.

Women selected for analysis in this study were born in that country because analysis for foreign born women is said to be more difficult due to uncertainty about their date of arrival in the country. There is no discussion about the quality of the LS, EDP or Norwegian Central Population Register and Educational Database.

Rendall et al. 2009 – *'Universal versus Economically Polarized Change in Age at First Birth: A French-British Comparison'*

This paper is similar in some ways to Ekert-Jaffé et al. (2002). The main hypothesis is that 'universalistic welfare regimes' reconcile conflicting demands of employment and motherhood, leading to a more universal age at first birth across socio-economic strata, while means-tested regimes produce increasingly heterogeneous distributions of age at first birth. The method used to assess this is analysis of the age at motherhood by pre-childbearing occupation across female birth cohorts ten years apart in the two countries. There is no discussion of data quality and methods used to improve the representativeness of the LS or EDP data to the population as a whole.

Portanti and Whitworth, 2009 – *'A comparison of the characteristics of childless women and mothers in the ONS Longitudinal Study'*

This research states that it is the first to have used the LS to explore lifelong childlessness. A specific cohort of women was selected - those born between 1956 and 1960 that were continuously resident in England and Wales during their childbearing years. The socio-economic characteristics of the women and their partnership status are considered and related to childlessness. The findings include that partnership status is the main factor associated with childlessness. Cohabiting women were found to be less likely than married women to be mothers. "Irrespective of their partnership status, women's own socio-economic characteristics, including economic activity and social class, are significantly associated with childlessness"

(Portanti and Whitworth, 2009 p.19). Beyond partnership status, socio-economic characteristics are still the main way of determining childlessness.

The merits of the large sample size of the LS are granted attention within the paper and it is highlighted that these allow for more robust statistical inference. However, within the limitations of the data there is a lack of discussion about the linkage of individuals between censuses, tracing of members and the linkage of births. The linkage of births to LS members and the degree to which this is a comparable sample with the overall population is an important issue. "Using the LS, we have been able to produce robust statistical results, as the LS is a large-scale nationally representative sample of women and their partners resident in England and Wales" (Portanti and Whitworth, 2009 p.18). The work compares the percentage of women remaining childless in national statistics against the figures in the LS and identifies that there are similar percentages. However, this does not mean that the LS is representative of women resident in England and Wales, or that there is no bias in the sample socio-economically (as suggested in the last section in the ONS reports on the linkage of LS members between census dates). Women in cohorts selected would be giving birth in the 1980s, a decade where there was generally a lower level of joining of births and the cumulative longitudinal linkage of LS members between census dates may impact on findings. The magnitude of the socio-economic trends identified could be because of the cumulative linkage of LS members between census dates.

Overall, although there has been a wealth of fertility research completed using the LS, there is a lack of recent research on births to LS members and the degree to which these accurately represent the population of England and Wales as a whole. The report by Babb and Hattersley (1992) is now somewhat dated, although the analysis completed is relevant for this current work to consider.

Within the academic literature there seems to be an understanding that there is a possible undercount of births in the LS compared to official statistics. In some of the research discussed in this section there has been recognition that there are general data quality problems with the LS which are a result of the way in which it functions. In particular, these problems are centred on the use of NHSCR data for embarkations from the LS. However, the ways in which this awareness has fed into data preparation and analysis are often unclear. Indeed, the representation of the LS sample that is selected is often not detailed in the published research. It seems common for work to acknowledge that there are weaknesses with the data but corrective action taken is not always transparent or explicitly clear. In the case of Ekert-Jaffé et al. (2002) there is a discussion of the EDP and data quality issues related to that, but not specifically for the LS. In contrast, other work like that of Rendall and Smallwood (2003) has taken very transparent action. This research used official statistics rates and the General Household Survey to improve the representativeness of the LS where there was a lower rate of linkage of births to sample members in the first two decades of the study. The secondary literature / reports from the ONS on tracing, sampling and linkage are often not discussed in published work. There is a need for research using the LS in fertility research to consider the residence of LS members and the representation of LS members and births to these members. The next section is concerned with hypothesising the main potential sources of error for this research to consider.

3.5 Conclusions

This chapter has outlined in precise detail the rationale which led to the development of the LS, the data of which it is composed and how it functions. It is important to be clear on all of these aspects before proceeding with research using the LS. It is also crucial to understand these aspects so as to be clear on where the LS has difficulties in representing the population of England and Wales accurately. Section 3.4 discussed how some fertility research using the LS has not discussed reports from the ONS on the quality of the dataset, nor been transparent on the selection of a sample for analysis. Yet given the statistics presented in Section 3.3 on the longitudinal linkage and tracing of LS members, it is necessary to consider this. It is essential for this research to address these issues before proceeding with fertility analysis using the LS, in order to be patently clear of the characteristics of the sample which is being analysed. This is especially true for work which is concerned with women who migrate to England and Wales.

To understand the data, it is important to understand the functioning of the LS. In Section 3.2 it was explained how persons become part of the LS and the ways in which they can enter and leave the study. In addition to these forms of entry and exit there is the joining of events data from the ONS Vital Statistics Output Branch and the NHSCR through updates to health records. However, the system for the capturing of migrations into and embarkations from England and Wales is still problematic, as individuals do not consistently embark in a recorded way through de-registration and therefore recording at the NHSCR. In
Section 3.3 this was explained in more detail, with the data sources of which the LS is composed and joining processes outlined. Quality aspects in this section are explained in terms of the 'tracing' and 'matching' of LS members and sampling fractions. Matching and tracing rates for the 'third decade' of the LS between 1991 and 2001 were the highest ever achieved, yet there are variations for different age groups; for women, the twenties showed the lowest rates of tracing. The primary concern is the way in which these LS members leave the study at some point in the intercensal period. Embarkations as recorded by the NHSCR are included in the LS data, and these can be used in analysis. In addition, the reports discuss 'attrition' in the data through unrecorded embarkations or not answering a census form.

Given these issues with the data, it is surprising that fertility research has largely ignored the published statistics on data quality. The last piece of research to have looked specifically at the strengths and weaknesses of the LS for fertility research is that by Babb and Hattersley (1992). This is now quite dated. Section 3.4 reviewed fertility research which has used the LS and the ways in which sample selection and data quality have been approached. Generally, research has taken one of three approaches. Some has ignored data quality dimensions and seems to assume that, by virtue of the large sample size, the LS is representative. A second set has recognised that the LS has some problems in terms of quality, does not take any action to explain incompleteness and, by definition, the sample used for analysis. A third strand of research has taken some form of corrective action. Rendall and Smallwood (2003) provide the most transparent explanation of their approach.

Using the Longitudinal Study for fertility research in England and Wales is viable, provided that limitations arising from its use are understood and appreciated. In England and Wales there is no other dataset that follows so many individuals over time, and thus makes the potential recording of changes among small population sub-groups possible. However, for such analysis to be undertaken, consideration of the nature of LS members continuously resident and accurately recorded is required. There is also no other dataset that functions by linking detailed event information related to 'events' like births, cancer registrations or deaths. Through such a dataset the classic concerns around accurate reporting in surveys and non-response in follow-up studies are eliminated. However, some research has not to discussed the degree to which there is the potential for missing cases and accurate representation, among some groups in the LS more so than others.

Through reviewing past fertility research using the LS, it is apparent that there has been no recent consideration of the sample women in the LS and their representativeness through time. In the next chapter, research questions which have emerged from this chapter on understanding the dataset will be outlined and answered, through the identification of different types of LS member. Through the creation of residence trajectories, the different forms of completeness and incompleteness in the dataset will become clear. These will allow the identification of a sample of women in the LS for analysis. Reports on how the LS functions and the review of linking and tracing provide important detail on the LS, but there is scope for expanding the understanding of this, especially with regard to use of the LS for fertility research.

Chapter 4

Creating trajectories for Office for National Statistics Longitudinal Study members and allocating exposure to risk of giving birth

Chapter abstract

Chapter 3 outlined in detail the way in which the Longitudinal Study (LS) combines routine administrative data to create individual records in the LS and highlighted that for the LS to be used with certainty, a rigorous approach to selecting a sample for demographic analysis is required. This chapter is concerned with coding the dataset based on the residence of individuals in the 1991-2001 and 2001-2007 periods. Through using 'residence trajectories', it is possible to identify LS members for whom there is complete information on residence between the 1991 and 2001 census dates and those where there is some form of incompleteness or error. The 'fourth decade' of the study is incomplete at present, but for the 2001-2007 period residence trajectories are created. Findings in this chapter provide foundations for selecting a sample for finer-grained fertility research using the LS than has previously been the case.

4.1 Introduction

The last chapter explained in detail the way in which the Office for National Statistics (ONS) Longitudinal Study (LS) combines data from various administrative sources within the ONS to create the full dataset. Given the various exit (or embarkation) points and the ways in which LS members can re-enter the study, the full complexity of the LS should be clear and careful attention must be paid to sample selection. This chapter focuses on selecting a sample from the LS for analysis. Establishing the exposure to risk for LS members in the 1991-2001 and 2001-2007 periods is the focus. This is done through the creation of idealised residence trajectories which show the residence patterns for LS members. Broadly, two types of LS member are identified – those who fall into a 'consistent' type and those who fall into an 'inconsistent' type.

The rationale for this work is simple. While in routine publications produced by the ONS after each LS-census linkage exercise there is background information on attrition in the LS, this information is often not used in the selection of a sample for analysis. Research using the LS often refers to the full LS sample when considering the representational aspects (i.e. sampling fractions), yet frequently selects a less dynamic sample for analysis (i.e. selects a sample of persons who have been consistently resident between census dates). Corroborating the typical selection criteria with reports from the ONS would suggest that this introduces an inherent bias in the sample selected for analysis. As a result of the work in this chapter, the number of LS members exposed to risk and the nature of their residence in the aforementioned time periods is clear. This enables a new approach to sample selection from the ONS LS.

4.2 Research questions

Chapter 3 explained some of the complexities with the way in which the ONS LS functions and the lack of detail on the data presented in fertility research using the LS. Past work which using the LS has not always detailed the suitability of the dataset or its potential use for fertility research. As a result of this, there are opportunities for research into the suitability of the LS for fertility research and specifically, for understanding women in the LS and their exposure to risk.

This section outlines the principal questions arising from the assessment of how the LS functions, analysis of reports by the ONS on data quality and reviews of published research with a focus on fertility using the LS. In order to use the LS for fertility research and to apply the methods from event history analysis, clarity on quality aspects is essential and the questions below have been devised with this in mind.

1. **How many female LS members have complete information on their residence throughout a decade, and are thus what could be called 'consistent cases'?**

   It is important to identify which female LS members were (according to available variables and event information) continually resident in the 1991-2001 period or, where there were events which meant that they left the study and were no longer included for a period of the decade, that period is known and accurately recorded. Variations by birth cohort will be of key interest.

2. **What characteristics are associated with incomplete information on the LS members and what events and information have incomplete information?**

   Through understanding where there is incomplete information for an LS member and what the incompleteness is, it is possible to correct the data we have for these individuals. Missing data might take one of the two following forms: (i) we know an event (embarkation or re-entry) occurred but do not know when, or (ii) we do not know if an event happened.

3. **How can we correct for missing cases in the dataset?**

   Based on some of the information that can be identified for incomplete cases and the information which has been gathered on complete cases (in 1), it will be possible to identify the best way of applying some form of correction to the data.

The next section outlines residence trajectories for LS members in the third decade of the LS and accounts for all LS members in this timeframe.

4.3 Creating and understanding trajectories for LS members in the third decade of the LS

4.3.1 Introduction

Post-census reports on the linkage of LS members from one census to the next, based on their characteristics, are a firm starting point, but accurate exposure-based analysis can be made only by understanding which LS members are present and when. The LS does not have a system of coding to allow the identification of which LS members are resident in a particular year. There is however information recorded on new entries to the study, embarkations from the study and re-entries. To aid with the understanding of potential sources of error in the LS, it is helpful to hypothesise and place numbers on individual 'trajectories'. By trajectory this work refers to life courses for LS members, including entry to the LS and exits from the LS. In hypothesising these, it is then possible to understand how the various data sources of which the LS is composed might lead to an accurate representation of the life course of an individual. This work attempts to understand how the LS captures residence in the 1991-2001 period. This is a staged process – firstly female LS members are considered purely in terms of exposure which is accurately recorded in the period between 1991 and 2001 (the third decade of the LS). These *consistent cases* also provide an initial idea of the numbers and percentages of the LS members for whom there is incomplete information on their whereabouts in the third decade. By 'consistent cases' this work is referring to those individuals where it is possible to be sure about their status as is recorded by the LS. The LS tells a coherent and complete story. It may not be complete in fact, but if it is incomplete this is because episodes are missing. It is thus assumed that there is complete information on the movements or residence in the time period; so as far as possible we can be sure that all embarkations and movements for those individuals have been recorded.

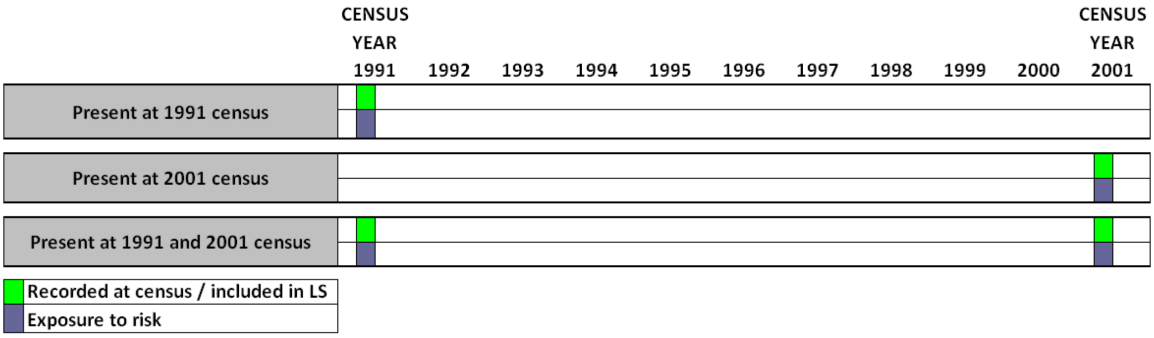
Inconsistent cases in the LS are the second broad category. Two forms of 'inconsistent case' can be identified: those where there was an event of some type (e.g. an unrecorded re-entry to the LS following an earlier, recorded embarkation) because other data tell us that this was the case (e.g. being resident at the 2001 census), and those where we cannot be sure if there was an event (e.g. non-residence at the census but no record of an embarkation). Analysis is completed by five-year birth cohort and enables age-based comparisons to be made. The census years of 1991 and 2001 are the beginning and end of the period of analysis, with the numbers of women resident at both being the initial point of interest. In the intervening period it is possible to identify how many individuals fit into the hypothesised trajectories developed, these being either consistent or inconsistent cases. Details on birth events to LS members are not of interest in this section.

Analysis is divided between LS members who are 'existing members', in that they have been part of the LS before the 1991 or 2001 census and 'new members' who are persons entering the LS for the first time between a census date.

4.3.2 A baseline of individuals resident at 1991 and / or 2001

Before explaining the different types of trajectory created and the numbers of LS members which fall into each, it is helpful to look at the number of LS members resident at a census in 1991 or 2001. These census dates are at either end of the third decade of the LS. Where an LS member was resident at a census, all the detailed socio-economic information which was collected on the census form is available. Figure 4.1 shows the possible residence at one of the census dates on either 21 April 1991 or 29 April 2001.

Figure 4.1: Baseline numbers of women in the LS at the 1991 and / or 2001 census



Source: James Robards, July 2010.

Table 4.1 shows the number of LS female members by birth cohort who were at the 1991 or 2001 census and also those who were at both. At both dates the total is fairly consistent and of those who were present in 1991, overall 80% (219,576) were at the 2001 census as well. From the cohorts from 1942 through to 1976 there is a slightly higher number of women than for the years before and after in all the categories. This is the first step in the analysis of residence for the 1991-2001 and 2001-2007 periods, as the census provides the baseline number of persons resident, is used for identifying individuals on the NHSCR and following their residence in England and Wales through time. Only those LS members who have been at a census have the full range of socio-economic variables collected through the census form. Although we can be sure that 80% of LS members who were at 1991 were also at 2001, we do not know what changes of residence these persons made between the two dates and thus their exposure to risk. This cross-sectional approach also means that any new entrants to the LS who are not at the census (i.e. persons who enter and embark or enter and die) are not included in analysis and the incorrect denominator or numerator may be used in analysis.

Table 4.1: Baseline numbers of women in the LS at the 1991 and / or 2001 census

BIRTH COHORT	Present at 1991 census	Present at 2001 census	Present at 1991 and 2001 census
1997 - 2001	-	12,283	0
1992 - 1996	-	16,062	0
1987 - 1991	14,952	17,741	13,107
1982 - 1986	16,625	16,295	14,508
1977 - 1981	16,098	15,782	13,459
1972 - 1976	16,483	16,800	13,674
1967 - 1971	19,761	20,425	16,794
1962 - 1966	21,276	21,515	18,489
1957 - 1961	19,929	19,673	17,651
1952 - 1956	18,168	17,806	16,228
1947 - 1951	19,507	19,005	17,549
1942 - 1946	17,707	17,052	15,831
1937 - 1941	14,526	13,825	12,785
Total	195,032	224,264	170,075

Own elaboration based on ONS LS, July 2010.

4.3.3 Hypothesising existing member consistent cases in the LS 1991-2001

As already described, the first step in this analysis is the identification of what can be termed consistent cases where there is complete, and it is assumed, accurate, information on the whereabouts of the LS member in the intercensal period. This subsection is concerned with 'existing member consistent cases' which are treated separately from 'new entrant consistent cases' (discussed in the next section). The distinguishing feature is that the LS members included in the residence patterns here were existing LS members as of the 1991 census. Figure 4.2 shows potential trajectories for the 11 types of existing member consistent cases which have been hypothesised.

LS members' resident at the 1991 census

Type 1 – An LS member who was resident at the 1991 and 2001 censuses, and has no recorded embarkation / migration or re-entry and no record of a death.

Type 2 – LS members who were resident at the 1991 census, yet died in the 1991-2001 period and were not resident in the 2001 census. They did not migrate or re-enter the LS in this period.

Type 3 – LS members who were resident at the 1991 census and who made an embarkation and re-entry which was then followed by a death. They were not at the 2001 census.

Type 4 – The LS member was resident at the 1991 census but then made a permanent embarkation and was not at the 2001 census.

Type 5 – The LS member was resident at the 1991 census, and there was an embarkation and re-entry in the decade. The person then embarked again and had not returned by the 2001 census.

Type 6 – The LS member was resident at the 1991 census, and there was an embarkation and re-entry in the decade. The person was present at the 2001 census.

LS members' not resident at the 1991 census

Type 7 – LS members who were not at the 1991 census, but re-entered in the 1991-2001 period and were at the 2001 census.

Type 8 – Cases where the LS member was not at the 1991 census, re-entered in the 1991-2001 period, and died at some point before the 2001 census. The LS member was not at the 2001 census.

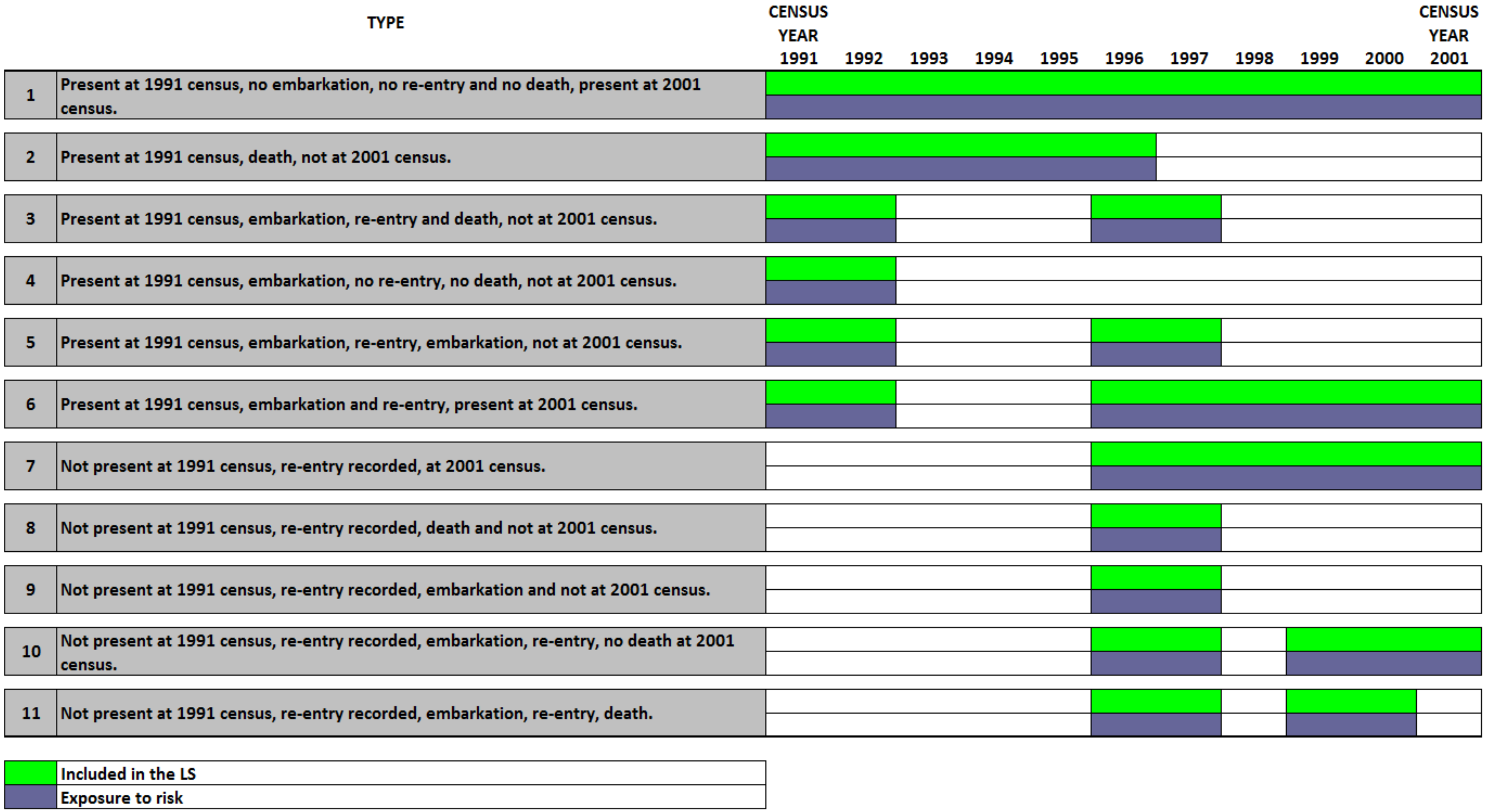
Type 9 – These are LS members not at the 1991 census, who re-entered the study in the 1991-2001 period and then embarked before the 2001 census. The LS member was not at the 2001 census.

Type 10 – Consistent cases where the LS member re-entered after the 1991 census, made another embarkation and then re-entered before being recorded at the 2001 census.

Type 11 – LS members who re-entered after the 1991 census, made an embarkation and re-entry before dying before the 2001 census.

These types are illustrated in the following figure, and numbers for each group given in 4.3.6.

Figure 4.2: Hypothesised consistent cases for existing members in the 1991-2001 period



Source: James Robards, July 2011.

4.3.4 Hypothesising new entrant consistent cases in the LS 1991-2001

In addition to the 11 consistent cases hypothesised and explained above, another six trajectories have been created for new entrants into the LS. The criterion is that the LS member entered the LS for the first time in the 1991-2001 period. These are discussed separately for clarity. Figure 4.3 shows the hypothesised trajectories for LS members who entered the LS in the third decade of the LS. Included among these new entrants are LS members who come from Scotland or Northern Ireland.

Type 12 – An LS member who entered the LS between 1991 and 2001 and was continually resident after entry (no recorded embarkation / migration or re-entry and no record of a death). This person was resident at the 2001 census.

Type 13 – LS members who entered between 1991 and 2001 and where there was no embarkation or re-entry, but where there was a death. The LS member was not at the 2001 census and so there is no socio-economic data for this person.

Type 14 – These are LS members who entered the LS in the 1991-2001 period. They then embarked and re-entered before 2001. They were recorded at the 2001 census.

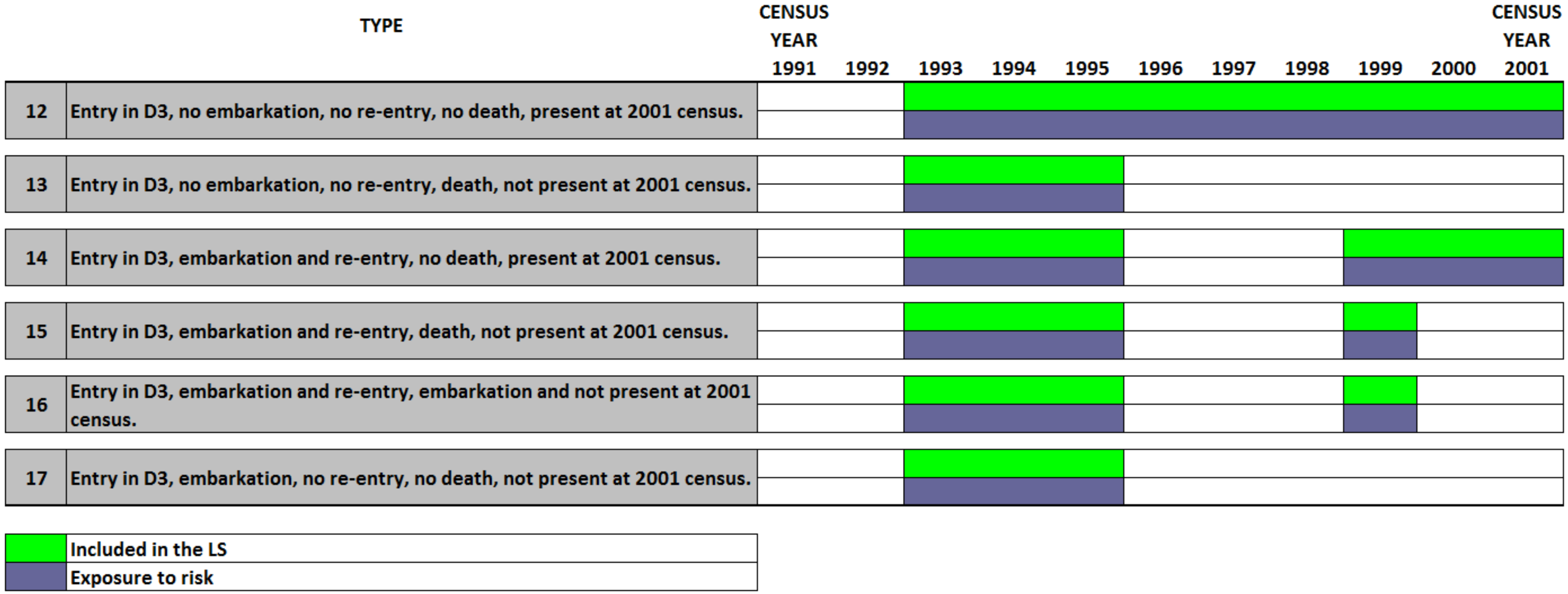
Type 15 – These are LS members who entered in the 1991-2001 period, then embarked and re-entered before dying. They were not at the 2001 census and thus there is no socio-economic data for them.

Type 16 – These are LS members who entered the LS in the 1991-2001 period and where there was an embarkation, re-entry and then another embarkation. There was no subsequent re-entry to the LS and they were not at the 2001 census.

Type 17 – These are members who entered the LS in the 1991-2001 period and where there was an embarkation. There was no subsequent re-entry to the LS and they were not at the 2001 census.

These types are illustrated in the following figure, and numbers for each group given in section 4.3.7.

Figure 4.3: Hypothesised consistent cases for new entrants into the LS in the 1991-2001 period



Source: James Robards, July 2010.

4.3.5 Results for consistent cases in the LS 1991-2001

Based on the hypothesised residence trajectories, it has been possible to classify types of LS members by birth cohort. With both the consistent and inconsistent cases in this analysis, there is no overlap between the broad types or the individual types; there is no double counting of LS members.

Table 4.2 shows the numbers for all consistent cases of new entrants and existing member types in the LS by birth cohort. Section 4.3.2 gave the raw number of traced women resident in 1991 and 2001 at the census – these figures are used as benchmarks for what LS members did in the third decade of the study. The census year numbers are simply a cross-section of LS members resident at a census and there is no condition that the LS member did not leave England and Wales in the third decade. Apart from any consideration of what LS members actually did in the third decade, the data gives a good indication of the relative exposure of LS members by different birth cohorts and age groups. An interesting point in this respect is the number of women who are in the LS consistent cases in the 1947-1966 birth cohorts. This constitutes the 'post-war baby boom' and has greater numbers than the other birth cohorts in the study.

4.3.6 Results for consistent existing member cases in the LS 1991-2001

Looking at each type in turn, it is clear that the largest number of all the consistent cases types is the first category - those persons where there was no recorded embarkation, re-entry or death in the third decade and the LS member was resident at both the census dates. This is the best possible outcome for understanding the exposure to risk for an LS member in the decade. By dividing the number of LS members of each birth cohort and Type (as in Table 4.2) which was at the 1991 census by the corresponding total number of LS members at the census (Table 4.1) it is possible to understand the percentage of LS members who were at the 1991 census and the trajectory that they fell into (e.g. for the 1987-1991 birth cohort in Type 1 we divide 13,071 (from Table 4.2) by 14,952 (from Table 4.1) giving 87.4%). Table 4.3 shows that in total, of cohorts born after 1937 87% of all LS members who were at 1991 fell into this type and 76% of LS members at 2001 were in this category.

However, these average figures conceal variations across age groups. These variations are best shown in Figure 4.4, which shows the percentage of Type 1 cases at the 1991 census. The drop in the proportion of Type 1 cases in the 1972-1976 and 1977-1982 birth cohorts is clear. Fewer of these members are resident compared with other birth cohorts; this suggests residential instability among those persons who would have been aged 20-29 years in 2001. Figure 4.4 also illustrates the way in which Type 1 contains most of the existing member consistent cases in the LS, as the total percentage of consistent cases is just above the line denoting type 1 cases. Table 4.4 shows that just 73.7% of LS members in the 1987-1991 cohort resident in 2001 were of Type 1.

LS members in Type 2 died during the third decade of the study and as would be expected are concentrated among older birth cohorts. Figure 4.5 shows the way in which the number of LS members in this category increase with age. As a percentage of all LS members who were at the 1991 census, the figures are very small. The precise date of death is known, but has not been used at this stage as it is sufficient here to know how many LS members die in the third decade.

There were no LS members in the third hypothesised trajectory in birth cohorts born after 1937. This hypothesised type was for LS members who embarked, re-entered and then died. Type 4 of the consistent cases was where LS members were at the 1991 census, then made an embarkation before 2001 and were not at the 2001 census. Table 4.2 shows that relatively small numbers of LS members fell into this type, and the largest numbers were for those in the 1957-1971 birth cohorts. In percentage terms, these LS members made up less than 1% of those LS members resident at the 1991 census. The small numbers who embarked and did not return pose some questions over embarkations from the LS and the number of members who return. Although it is hard to be sure, there is the chance that the person who embarked may have actually returned, but not rejoined the LS. This would have happened if the LS member embarked and returned, but did not register with a GP, and was recorded as not having an NHSCR record. It is also possible that the LS member embarked and returned, registered with a GP and re-joined the LS, but was not included because they were not captured at the 2001 census, despite being resident at that time.

Types 3, 5, 8,10 and 11 contain small numbers or no cases (where the type is not included in the table there were fewer than 10 cases in total). Type 6 cases are those LS members who were at the 1991 census and then embarked and re-entered at some point before the 2001 census. Types 8 and 9 are the most problematic consistent types because these are cases where persons re-entered the LS in the third decade having been absent in 1991 but then embarked before the 2001 census. This means that there is a lack of (at least recent) socio-economic information for these individuals. In the birth cohorts since 1937, just 271 members fell into Type 9 and 27 were in Type 8.

4.3.7 Results for consistent new entrant cases in the LS 1991-2001

The focus of this thesis is migrants and therefore new entrants to the LS are treated separately so the number in each trajectory can be clearly distinguished and residence at a census understood. Residence at a census is important because of the range of socio-economic variables collected from the census form which will be crucial for subsequent analysis. In Table 4.2 the right hand panel provides the number of new entrants recorded by the LS and the types into which they fall. As with Type 1 for the existing members, Type 12 for new entrants is the best scenario with the member entering the study in the third decade and then being resident until at least the 2001 census with no recorded embarkation or re-entry to the study. By birth cohort there is clearly a large number of members who are aged 20-30 years (as of the 2001 census) who enter in the third decade of the study. In total, 4,034 entries are recorded in the 1967-1981 birth cohorts. Consistent new entrant Types 13 and 14 contain small numbers and Types 15 and 16 do not have any in cases in the 1937-2001 birth cohorts.

Type 17 contains new entrants to England and Wales who entered the LS in the 1991-2001 period and where there was an embarkation. There was no subsequent re-entry to the LS and the LS member was not at the 2001 census. Just over 1,000 members are in this category and most of these are in the 1967-1976 birth cohorts. Table 4.4 shows that LS members in Type 12 make a contribution to the total at the 2001 census, particularly in the 1967-1971 cohort (6.4%), 1972-1976 (9.7%) and 1977-1981 (7%). As a percentage of some birth cohorts at the 2001 census, there is a significant contribution of new entrants into the LS. The greatest number of entrants to the LS is from the 1967-1981 birth cohorts, and in particular the 1972-1976 birth cohort.

4.3.8 Overall results for all consistent cases in the LS 1991-2001

Looking at the overall picture of consistent cases in the LS, it is clear that it is those aged 20-30 years in 2001 where there are some of the most interesting trends. This age group is one of the key childbearing ages. Figure 4.6 shows that the percentage of LS members in the birth cohorts 1972-1981 resident in 2001 who were not new entrant consistent cases is smaller than for other cohorts. At the same time, these are the peak birth cohorts for new entrants / immigrants into the LS – Figure 4.6 shows the way in which the different categories compare. There are clearly fewer LS members in the consistent types for the most recent birth cohorts. The lowest representation is in the 1972-1976 cohort, where there are just 81.5%, making up consistent cases in the existing member group (except the 1987-1991 birth cohort which overlaps with the 1991 census in April 1991). Following this, there is substantial rise back towards 1991. This highlights that there is a decline in the ability of the LS to accurately follow LS members in these cohorts for the third decade. Through the understanding of the inconsistent cases, a picture of the impact of this should become clearer. This finding corresponds broadly with the analysis made in ONS reports.

The next section discusses the inconsistent cases in the third decade, and through understanding the numbers in this category it is then possible to be clearer on who the LS can accurately represent for exposure-based analysis.

Table 4.2: Numbers of LS members falling into consistent cases by migration status 1991-2001

Migrant status	*Existing members*									*New entrants*					All
BIRTH COHORT	Type 1	Type 2	Type 4	Type 6	Type 7	Type 8	Type 9	Type 10	*Total*	Type 12	Type 13	Type 14	Type 17	*Total*	TOTAL
1997 - 2001	*	*	*	*	*	*	*	*	*	209	*	*	11	220	220
1992 - 1996	*	*	*	*	*	*	7	*	13	470	*	*	65	538	551
1987 - 1991	13,071	15	67	17	11	*	15	*	13,196	547	*	*	54	606	13,802
1982 - 1986	14,477	22	49	23	35	*	11	*	14,617	457	*	*	23	483	15,100
1977 - 1981	13,415	34	46	24	37	*	16	*	13,572	1,105	*	*	88	1,201	14,773
1972 - 1976	13,621	49	52	35	43	*	33	*	13,833	1,630	*	*	281	1,919	15,752
1967 - 1971	16,691	86	146	66	45	*	42	*	17,076	1,299	*	18	252	1,575	18,651
1962 - 1966	18,400	102	151	50	46	*	56	*	18,805	861	*	10	124	995	19,800
1957 - 1961	17,595	168	122	30	51	*	33	*	17,999	478	*	*	76	562	18,561
1952 - 1956	16,178	233	88	27	31	*	16	*	16,573	299	*	*	32	335	16,908
1947 - 1951	17,523	364	64	17	30	*	13	*	18,011	199	*	*	32	231	18,242
1942 - 1946	15,806	551	43	13	20	*	12	*	16,445	123	*	*	22	148	16,593
1937 - 1941	12,773	634	42	*	25	*	17	*	13,497	78	*	*	25	110	13,607
Total	169,550	2,258	870	308	380	27	271	16	173,680	7,755	27	64	1,085	8,931	182,611

**Numbers deleted to allow clearance from Micro-data Analysis and User Support (MAUS), ONS.*

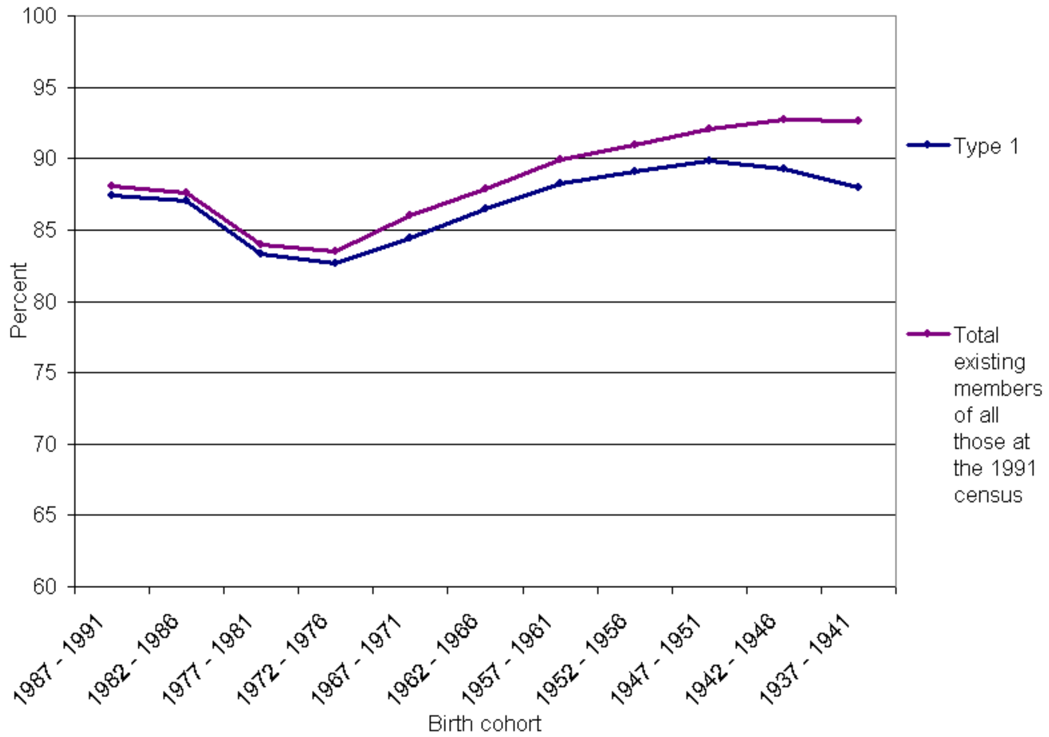
Source: ONS LS / James Robards, July 2010.

Table 4.3: Percentage of LS members by type, resident at the 1991 census

Existing members					
BIRTH COHORT	Type 1	Type 2	Type 4	Type 6	*Total*
1987 - 1991	87.4	0.1	0.4	0.1	88.1
1982 - 1986	87.1	0.1	0.3	0.1	87.6
1977 - 1981	83.3	0.2	0.3	0.1	84.0
1972 - 1976	82.6	0.3	0.3	0.2	83.5
1967 - 1971	84.5	0.4	0.7	0.3	86.0
1962 - 1966	86.5	0.5	0.7	0.2	87.9
1957 - 1961	88.3	0.8	0.6	0.2	89.9
1952 - 1956	89.0	1.3	0.5	0.1	91.0
1947 - 1951	89.8	1.9	0.3	0.1	92.1
1942 - 1946	89.3	3.1	0.2	0.1	92.7
1937 - 1941	87.9	4.4	0.3	0.0	92.6
Total	86.9	1.2	0.4	0.2	88.7

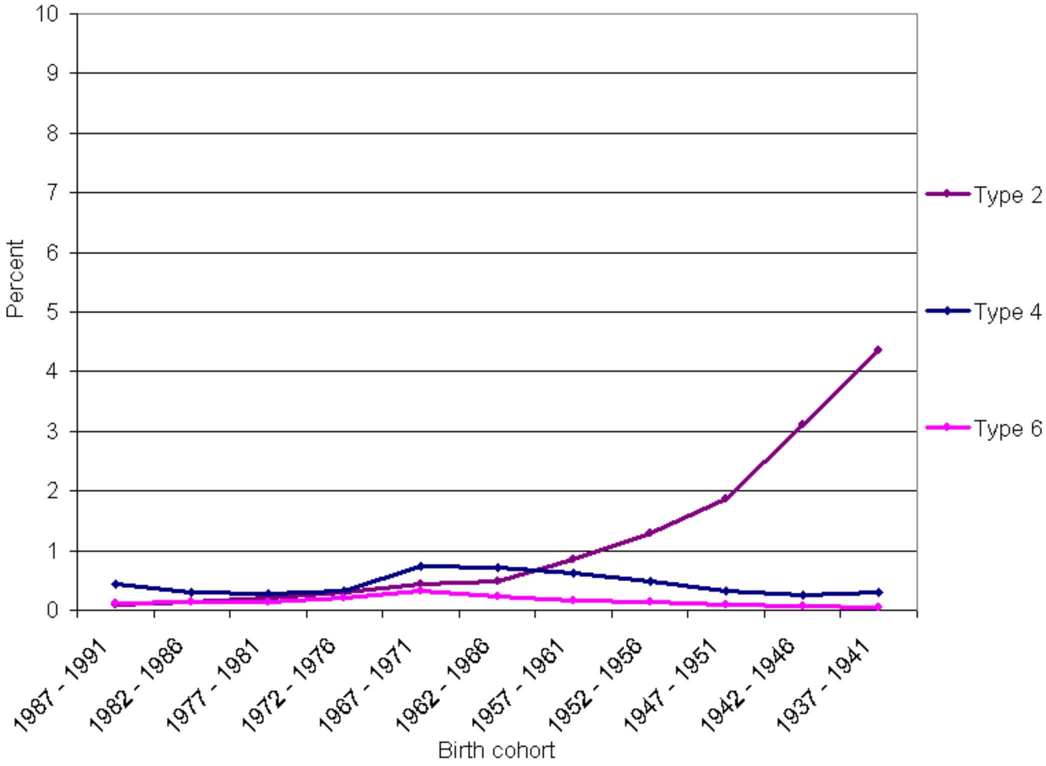
Own elaboration based on ONS LS, July 2010.

Figure 4.4: Percentage of existing consistent LS members by type, resident at the 1991 census



Own elaboration based on ONS LS, July 2010.

Figure 4.5: Percentage of existing consistent LS members by type, resident at the 1991 census



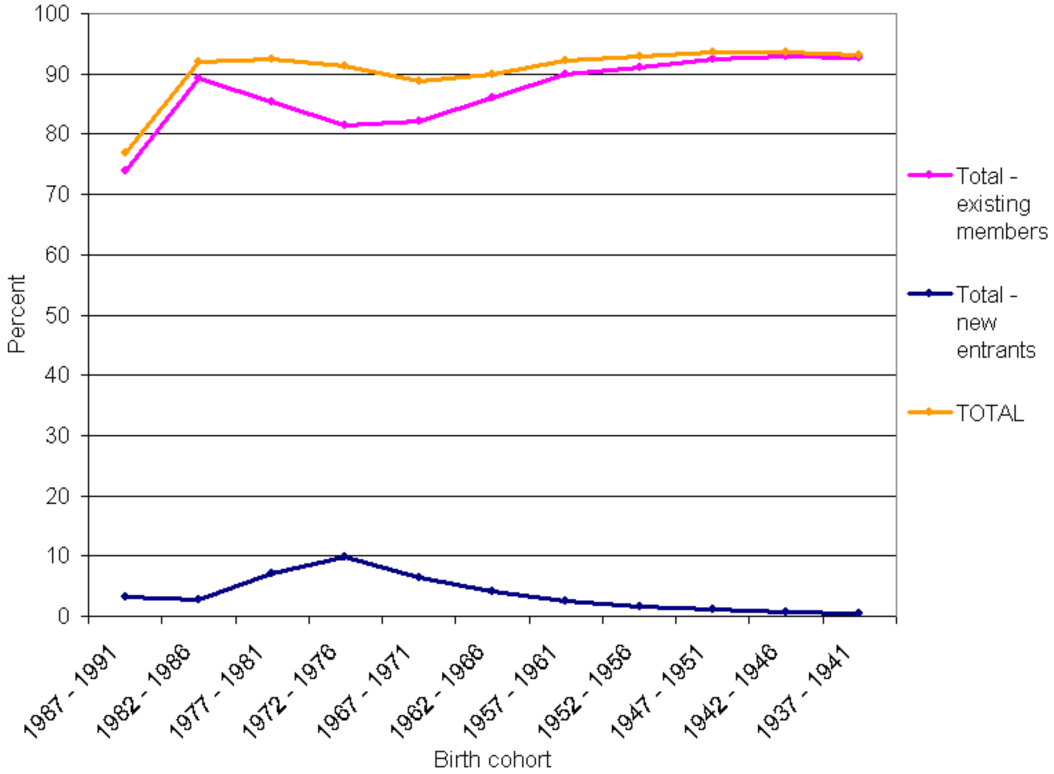
Own elaboration based on ONS LS, July 2010.

Table 4.4: Percentage of LS members by type, resident at the 2001 census

	Existing members				*New entrants*			All
BIRTH COHORT	Type 1	Type 6	Type 7	*Total*	Type 12	Type 14	*Total*	TOTAL
1997 - 2001	-	-	-	-	1.7	*	1.7	1.7
1992 - 1996	-	-	0.0	0.0	2.9	0.0	2.9	3.0
1987 - 1991	73.7	0.1	0.1	73.8	3.1	0.0	3.1	76.9
1982 - 1986	88.8	0.1	0.2	89.2	2.8	0.0	2.8	92.0
1977 - 1981	85.0	0.2	0.2	85.4	7.0	0.1	7.1	92.4
1972 - 1976	81.1	0.2	0.3	81.5	9.7	0.0	9.8	91.3
1967 - 1971	81.7	0.3	0.2	82.3	6.4	0.1	6.4	88.7
1962 - 1966	85.5	0.2	0.2	86.0	4.0	0.0	4.0	90.0
1957 - 1961	89.4	0.2	0.3	89.8	2.4	0.0	2.5	92.3
1952 - 1956	90.9	0.2	0.2	91.2	1.7	*	1.7	92.9
1947 - 1951	92.2	0.1	0.2	92.4	1.0	*	1.0	93.5
1942 - 1946	92.7	0.1	0.1	92.9	0.7	*	0.7	93.6
1937 - 1941	92.4	0.0	0.2	92.6	0.6	*	0.6	93.2
Total	75.6	0.1	0.2	75.9	3.5	0.0	3.5	79.4

Own elaboration based on ONS LS, July 2010.

Figure 4.6: Percentage of LS members resident at the 2001 census who were consistent cases



Own elaboration based on ONS LS, July 2010.

4.3.9 Hypothesising existing member inconsistent cases in the LS 1991-2001

Inconsistent cases in the LS are more complex than the consistent cases which have been outlined. Broadly, there are two types of inconsistent cases in the data:

1. Records for LS members where it is known that an event happened but the date is unknown (e.g. an unrecorded re-entry to England and Wales but not the LS; known about through residence at the next census).
2. LS member records where it is not possible to be sure that an event actually happened (e.g. a person not being recorded at the 2001 census yet there not being any record of his or her death).

In considering the types of inconsistent cases in the LS in the 1991-2001 timeframe, it is helpful to split the impact of the incompleteness into two forms. There are LS members where we might assume that a person is resident, yet they are in fact not and the denominator (women at risk of birth) is *inflated*, and cases where the LS does not include someone when it should and is therefore *deflating* the denominator. In the case of the error where the denominator would be inflated, this might be the result of attempting to record persons who have made an unrecorded embarkation. For the deflationary impact, an example of this could be where the LS member was resident in England and Wales yet had not registered with a GP and triggered entry to the LS, despite having been born on one of the LS dates and therefore being an 'LS candidate'.

In hypothesising the potential inconsistent cases in the LS, various Types were created, some of which have been discarded in the final version outlined here. The permutations possible for hypothesised trajectories are numerous, so it is necessary to consider the most pertinent forms of incompleteness in the data, namely unrecorded embarkations and re-entries. As already mentioned, the complexity of the LS and the various flows into and out of the data are numerous.

Figure 4.7 shows the final hypothesised inconsistent cases on which data has been collected.

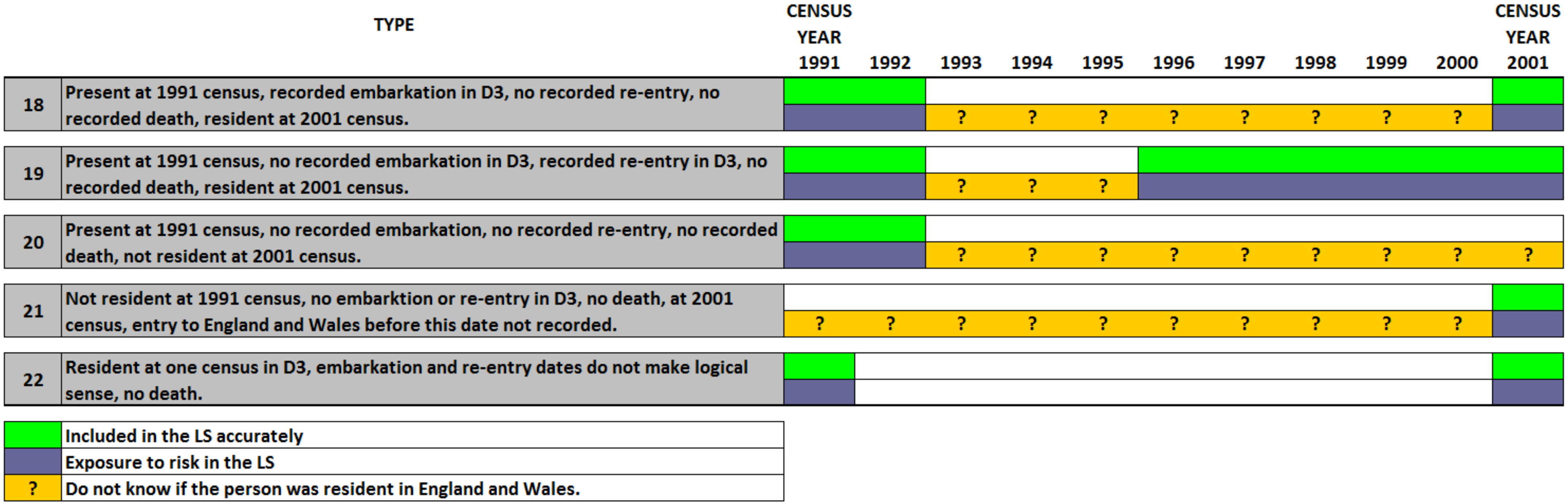
> **Type 18 –** LS members who was present in 1991 and who had a recorded embarkation in the third decade of the LS, but no recorded re-entry and then the person was identified at the 2001 census. The re-entry to England and Wales was not recorded, and by taking the 1991 and 2001 cross-sections (ignoring embarkations) it would be possible to give the LS member residence for the entire timeframe without taking into account the unrecorded re-entry.
>
> **Type 19 –** These are LS members who made an unrecorded embarkation, but a recorded re-entry. Effectively these are the opposite of Type 18. If these members are assumed to be resident for the whole decade of interest then the denominator would be inflated by these members, as they are thought to be resident but have left.

Type 20 – This is the most problematic case, as will be explained. Although the person was at the 1991 census, they were not at the 2001 census and there is no record of an embarkation, re-entry or death. This means that, apart from any event data there might be for that person, there is a scarcity of information on that LS member's exposure to risk in the 1991-2001 period. It is possible that the LS member was resident in England and Wales in 2001, yet for some reason was not recorded on a census form.

Type 21 – These are LS members who were not resident in 1991 but who may have been resident in the LS at some point in the past. This Type also includes LS members for whom this is the first census and the first time they are included in the LS. They 'appeared' at the 2001 census and it is not known when they may have entered England and Wales before this date (unless they were an LS member in the pre-1991 period).

Type 22 – This is a case where the LS member was at the 1991 census, but where there is an inconsistency in the recorded embarkation and re-entry data. The dates are incorrectly recorded, or there is an inconsistency where the date of re-entry was before the embarkation. This means that an embarkation or re-entry may have been missed.

Figure 4.7: Hypothesised inconsistent cases for existing members in the 1991-2001 period



Source: James Robards, July 2010.

4.3.10 Hypothesising new entrant inconsistent cases in the LS 1991-2001

In addition to the existing member cases, there are many forms of incompleteness for the new entrants to the LS in the 1991-2001 period. These are explained individually:

Type 23 – This is similar to existing member Type 20. In this case the person enters the LS and then goes 'missing' before 2001, with no record of a death or embarkation. The date at which this happens is not known, although the use of event information (i.e. births or cancer registrations) offers the potential to identify for how long LS members are resident after their entry.

Type 24 – These are LS members where there is a recorded death and yet the LS member is found to be at the 2001 census.

Type 25 – These are LS members where the entry to the LS is recorded, and where there is an embarkation and re-entry. At some later point the LS member leaves as there is no recorded death or embarkation, or the LS member is not recorded at the 2001 census. There was no recorded embarkation after the re-entry to the LS, so either the LS member embarked again without it being recorded or they were unrecorded at the 2001 census.

Type 26 – These are LS members who make a recorded embarkation but are recorded at the 2001 census. The date of re-entry to England and Wales is not known.

Type 27 – This is a case which is similar to Type 22 under the existing member cases. The LS member was recorded as entering, and then there is an inconsistency in the embarkation and re-entry dates within the third decade. The LS member is at the 2001 census.

Type 28 – LS members who are new entrants where there was an entry date, no recorded embarkation and then a re-entry date. This means that they were ‘missing’ for an unknown period from the date of entry, through to the date of re-entry.

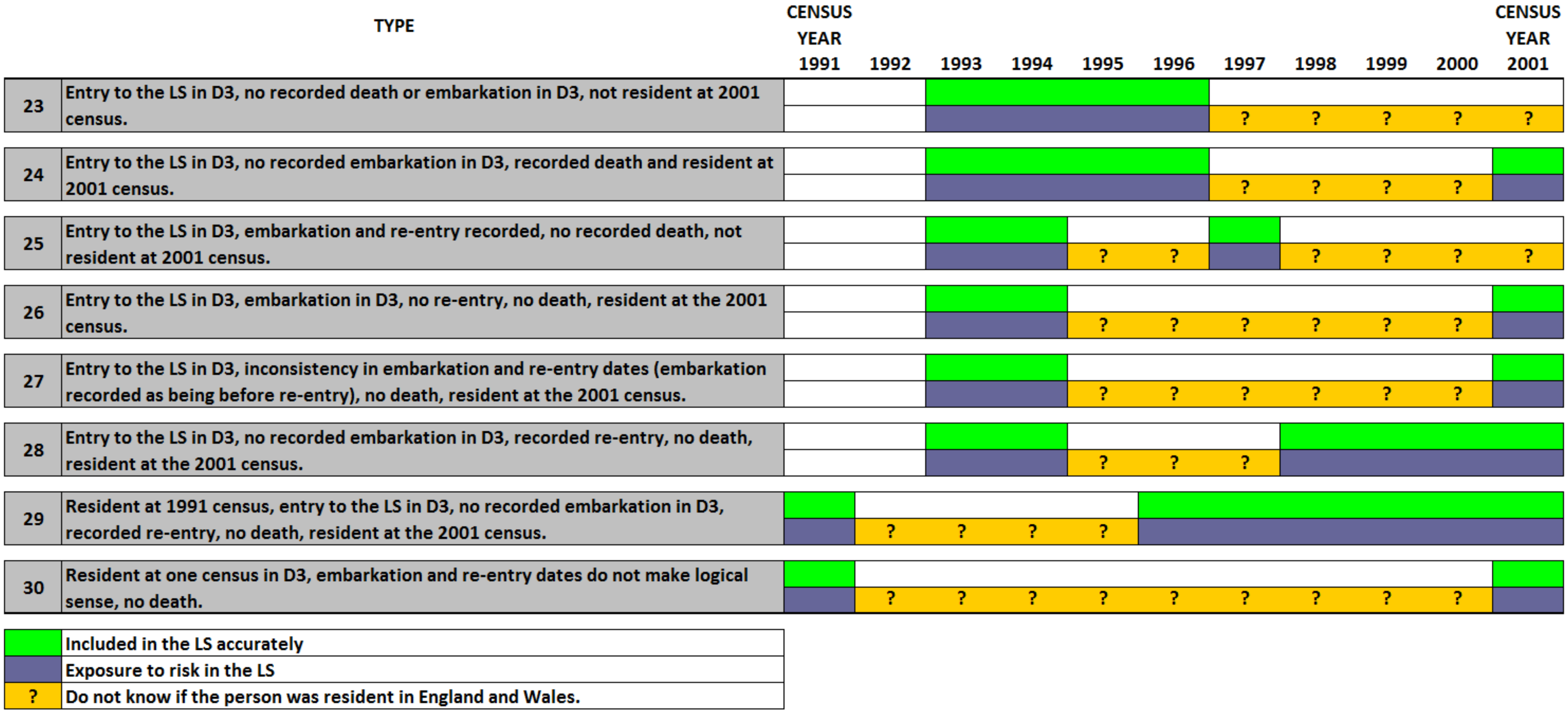
Type 29 – LS members who were recorded as entering the study in the third decade, yet were also recorded as being at the 1991 census. As with Type 26, there must be a recorded re-entry but no embarkation, and then residence at the 2001 census.

Type 30 – These are persons who are resident at one of the census dates in the third decade but where the embarkation and re-entry dates do not make logical sense, in that the embarkation is recorded as being after the re-entry.

To consider the relative importance of each of these hypothesised trajectories, syntax was devised in SPSS to create new variables required to calculate with the date information provided and understand the residence of the LS members. The next section outlines the numbers of LS members falling into each of these types, and also the number in relation to the birth cohort and the sample at the 1991 / 2001 census dates.

Presentation of results for the inconsistent cases is more complex than for the consistent cases. There are more possible residence trajectories and, as will become apparent, there are small numbers for many of the trajectories that have been hypothesised. Each member in the LS has coding; this is important for onward analysis using the LS.

Figure 4.8: Hypothesised inconsistent cases for new entrants in the 1991-2001 period



Source: James Robards, July 2010.

4.3.11 Results for inconsistent existing member cases in the LS 1991-2001

Table 4.5 shows the number of LS members who fall into one of the inconsistent types. The first form of inconsistency – where there was a recorded embarkation and then no recorded re-entry before 2001, yet recorded residence at the 2001 census (Type 18) – has relatively small numbers. Only 26 cases in the LS fall into this type. There are small numbers for each birth cohort. This trajectory shows that there are very small numbers of LS members where there is a recorded embarkation and no recorded re-entry, yet the person is at the next census. Given the requirement that a person must re-register with a GP to re-enter the study, one would expect that perhaps there would be more people in this category. This suggests that individuals re-register with their GP on arrival back in England and Wales. It is possible that there are LS members within the consistent cases, where it is thought that the person is continuously resident but there was actually an embarkation and re-entry which was not recorded. It is simply not possible to identify any such case as this because there is no additional information which can be used. This means that we have to assume that the LS case is consistent; there is no additional information to suggest otherwise.

Importantly, Type 18 should be understood in the context of the number of cases where there was an unrecorded embarkation, Type 20 in Table 4.5. These are cases where the LS member was at the 1991 census and then embarked at some point after 1991 but this was not recorded. There are 21,382 cases (in birth cohorts 1937-1991) which fall into this type. There are distinct age group trends and these are shown over the next few pages in full. Type 20 might also arise from the following situations:

- An unrecorded embarkation - of the potential reasons, this is one of the most likely, given the problems with the LS in recording embarkations.
- Not being recorded at the 2001 census - it is possible that the LS member was in England and Wales on the census night but was not recorded on a census form and therefore not recorded as being resident (a Type 1 case). There is no way of knowing if this was the situation.
- An unrecorded death - this is unlikely, given the way in which this is one of the events that is ‘double recorded’, as explained in Chapter 3.

Comparing the numbers in Types 18 and 20 suggests that although embarkations from England and Wales might not always be recorded, the LS is good at recording re-entries to England and Wales.

Table 4.5: Numbers of LS members of an inconsistent type by migration status 1991-2001

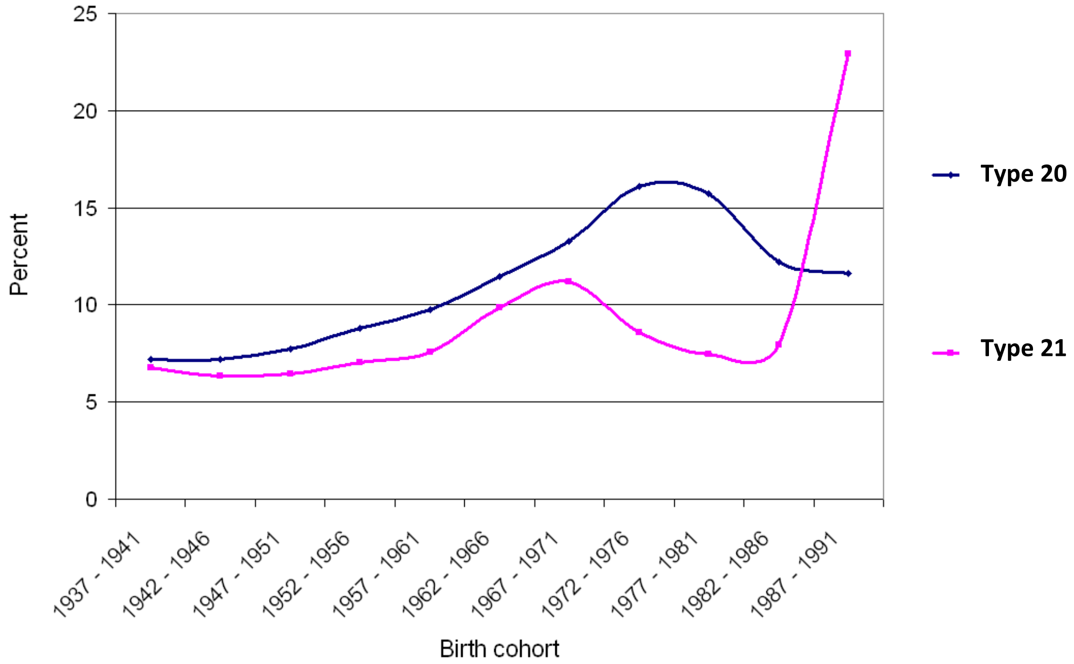
	Existing members			*New entrants*		*All*
BIRTH COHORT	Type 20	Type 21	Total	Type 23	Total	TOTAL
2007 - 2011	*	*	*	*	*	*
2002 - 2006	*	*	*	*	*	*
1997 - 2001	*	*	12,075	218	218	12,293
1992 - 1996	*	*	15,582	550	557	16,139
1987 - 1991	1,735	4,063	5,822	674	682	6,504
1982 - 1986	2,027	1,293	3,333	618	621	3,954
1977 - 1981	2,535	1,179	3,731	2,079	2,087	5,818
1972 - 1976	2,653	1,442	4,123	3,628	3,650	7,773
1967 - 1971	2,625	2,283	4,951	2,664	2,688	7,639
1962 - 1966	2,437	2,117	4,589	1,463	1,471	6,060
1957 - 1961	1,941	1,491	3,462	871	877	4,339
1952 - 1956	1,598	1,256	2,866	491	493	3,359
1947 - 1951	1,512	1,223	2,752	311	313	3,065
1942 - 1946	1,271	1,078	2,357	290	293	2,650
1937 - 1941	1,048	937	1,991	253	255	2,246
Total	21,382	18,362	39,997	14,110	14,205	54,202

Own elaboration based on ONS LS, July 2010.

For Type 21, unexpected arrivals at the 2001 census, there are 18,362 cases. Values for the 1992-2001 period are not shown as these include new births into the LS since the 1991 census. The higher figure for the 1987-1991 cohort is likely to arise from the number of LS members entering the study through being born on an LS date in the remainder of 1991 after the census. Overall based on the numbers of unrecorded entries to the LS in 2001 and the number of LS members who leave in an unrecorded way after 1991, it is clear that the LS is better at capturing new members to England and Wales than recording embarkations. This is consistent with the findings in Chapter 3 using the NHSCR GP register numbers in relation to the ONS mid-year estimates (in summary there is an over-count of persons on the NHSCR when compared to the mid-year estimates).

Figure 4.9 shows the percentage of LS members at the 1991 census who make some form of unrecorded embarkation between 1991 and 2001. This is highest in the key birth cohorts of 1972-1981. The figure also shows the percentage of LS members who were at the 2001 census for the first time (compared with all who were at 2001) and did not enter the LS as immigrants before this date. Notable is the rise in the 1987-1991 birth cohort which in part is because all births in 1991 after the census are included.

Figure 4.9: Percentage of LS members of Type 20 (unrecorded embarkation) as a percentage of those resident at the 1991 census and Type 21 (unrecorded entry) as a percentage of those resident at the 2001 census



Own elaboration based on ONS LS, July 2010.

Type 22, cases where there is an inconsistency between the embarkation and re-entry information and residence at one census was recorded, consists of relatively small numbers. These cases form what could be termed a ‘residual category’, where any further detailed decomposition would bring minimal returns and the level of incompleteness and inconsistency is too high.

4.3.12 Results for inconsistent new entrant cases in the LS 1991-2001

Table 4.5 showed the number of new entrants of different types in the LS. In the inconsistent types the most numerous case is Type 23, where there is a recorded entry into the LS and then an unrecorded embarkation at some point before 2001. The trends in this type mirror those of Type 20 for the non-immigrants – the 1972-1976 birth cohort has the highest number of unrecorded embarkations. For the other trajectories there are relatively small numbers. This shows that the LS is losing new entrant members between their entry to the LS and the next census, or did not record new entrants at the 2001 census. However, as was shown in Figure 4.6, the immigrants into the LS push the overall percentage of consistent cases up above 90% for the birth cohorts where there are the largest losses. So in the age groups where there are the greatest rates of attrition between one census and the next, are those where there are the highest rates of new entrants joining the LS. This highlights that the 1972-1976 birth cohort is well represented, even with some losses, and that using residence at 1991 and 2001 census as a criteria would exclude the large number of new entrants to the LS in the third decade.

4.3.13 Overall results for all inconsistent cases in the LS 1991-2001

While there are large numbers of LS members falling into the inconsistent types for the 1991-2001 period, there are fewer forms of residence types for these members. The most numerous types are those where there has been an unrecorded embarkation, or the person has not appeared on a census form in 2001. By age group, women in their twenties are most likely to fall into this category. There are also high numbers of women who enter at some point between 1991 and 2001 but the date of entry is not recorded.

Overall, the message from the analysis of inconsistent cases is that the LS is a better dataset at capturing entrants to the study than it is at recording embarkations. This is likely to be a result of the way the study operates – while new entrants are recorded through registration with a GP, embarkations are recorded in a less precise way. In Chapter 3 the analysis of NHSCR data on the degree to which the patient register is over-inflated compared with mid-year estimates highlighted that the LS is likely to be of a similar form.

4.3.14 Conclusions for trajectories in the third decade of the LS

Although in some respects now a little dated, the 1991-2001 period is the last complete decade of the LS. This is important to understand, as it gives two datum points, 1991 and 2001, which can be used to understand the flow of members into and out of the study in the period between. The identification of consistent and inconsistent cases is a first step in analysis which will fully understand the study members in the LS who are exposed to risk in a particular year. Reporting by the ONS on attrition from the LS is helpful, but the dataset is not coded in a way that makes the cases identifiable. This is problematic for analysis by single year in the periods between census dates. For understanding the exposure to risk for LS members and the calculation of fertility rates, the approach taken allows the identification of a sample which is more robust and well understood than is conventionally the case in analysis using the LS.

Overall, a high percentage of LS members in 1991 were present in 2001. However, there are noticeable difference for the birth cohorts used. Women in the 1976-1982 birth cohort were least likely to be linked between the two dates. The date of and precise reason for attrition are unknown; this may be a result of not completing a census form or that the person left England and Wales in the third decade and was not resident in 2001.

The results for consistent cases show that the LS has a high proportion of cases which can be used with complete confidence about the whereabouts of the LS member. However, when the results are assessed by age group problems with the accuracy of the LS for understanding the exposure to risk, and by definition fertility, for selected age groups become apparent. For some age groups we know where around 90% of LS members at the 1991 census were between then and 2001, but for the more recent birth cohorts there are fewer members that are resident continuously between the two census dates without any form of embarkation. In particular, the 1972-1981 birth cohort has the highest percentage of LS members resident in 1991 who make an unrecorded embarkation or are not recorded at the census. These members would be aged between 10 and 19 years in 1991 and 20 and 29 years in 2001. The 1972-

1976 birth cohort has high numbers of members which entered the study at some point after the 1991 census. Although this cohort has lower numbers of LS members continuously resident between the two census dates, it also has a higher number of women entering in these years. This adds further information on the sample for analysis and through looking at the entries to the LS and the numbers of people continuously resident, it becomes apparent that there is a high degree of flux in terms of the movements of the individuals in this age group into and out of England and Wales. It was illustrated that, overall, the 1972-1976 birth cohort is that with the lowest percentage of consistent cases based on the typologies that were devised. This also serves to highlight that using only the population resident at both the 1991 and 2001 census dates for analysis means that many women in the key childbearing ages would be excluded from analysis, and also that new immigrants to England and Wales in these age groups would be excluded.

When syntax was run for the inconsistent cases in the LS, many of the trajectories created had small numbers. The largest type in the data was for those where the LS member was at the 1991 census, not at the 2001 census, and where there was no recorded embarkation, re-entry or death. The point at which these persons left the study is not known. It could be that they were resident at the 2001 census but not recorded or identified there, or they may have left at an earlier date.

From the hypothesising of types of consistent and inconsistent cases, it has become apparent where there are potential problems with the LS in understanding the exposure to risk of LS members. Accounting for the residence of persons recorded at the 1991 census between then and the 2001 census is a crucial step before further analysis is made in calculating the births to LS members and the number of LS members resident. It is also crucial before proceeding with more advanced exposure and event-based techniques for the analysis of changing fertility in England and Wales. The next step in this analysis will be to consider the trajectories for those LS members who are resident in the fourth decade of the study after the 2001 census. This includes those who were resident in 2001, those LS members who re-enter the study and new entrants to the LS after the 2001 census. Without complete information on the residence for the fourth decade (linkage of the 2011 census data is on-going), there is no right censoring point for LS members. Having an identifier in the dataset for onward analysis is an important step for fertility analysis post-2001.

4.4 Creating and understanding trajectories for LS members in the fourth decade of the LS

4.4.1 Introduction

The third decade of the LS is just one of the 'complete' decades of the study. At the current time, work is underway on the LS-census link. There is no way of knowing which of the LS members currently in the study will be at the 2011 census. The results on who was at the 2001 census give an indication of who we would expect to be at 2011, so the analysis of the 1991-2001 period is very useful for considering the post-2001 picture. As with the third decade, it is possible to hypothesise and investigate the numbers of LS members who fall into different types of residence trajectories. Without an end date (a census) which can be used in the analysis, there is no firm 'right censoring' to the time period, the latest date for which data is available is 2007 and it is this year which is used in this analysis.

4.4.2 Hypothesising existing member consistent cases in the LS 1991-2007

As with the trajectories for the third decade of the LS, it is possible to hypothesise trajectories for consistent cases in the LS in the 2001-2007 period. Figure 4.10 shows trajectories for LS members in the 2001-2007 period. These are 'existing member' cases; new entrants to the LS in the time period are considered separately in the next section. There are nine types of consistent case hypothesised. The difference from the third decade data is that there is no right censoring in the form of the census.

> **Type 1 –** An LS member who was resident at the 2001 census and has been continually resident in the 2001-2007 period. (There has been no record of an embarkation or re-entry and no record of a death).

Type 2 – LS members who were resident at the 2001 census and died in the 2001-2007 period. They did not migrate or re-enter the LS in this period.

Type 3 – These are LS members who were resident at the 2001 census, where there was an embarkation and re-entry which was then followed by a death.

Type 4 – This type is where the LS member was resident at the 2001 census, made an embarkation and re-entry.

Type 5 – Cases where the member was at the 2001 census and then made an embarkation with no recorded re-entry as of 2007.

Type 6 – These are persons who re-entered the study after the 2001 census and where there has been no subsequent embarkation or death.

Type 7 – Cases where the LS member re-entered after the 2001 census, did not embark again and died.

Type 8 – LS members not at the 2001 census who re-entered the study after the census, then embarked and re-entered again and were resident in 2007.

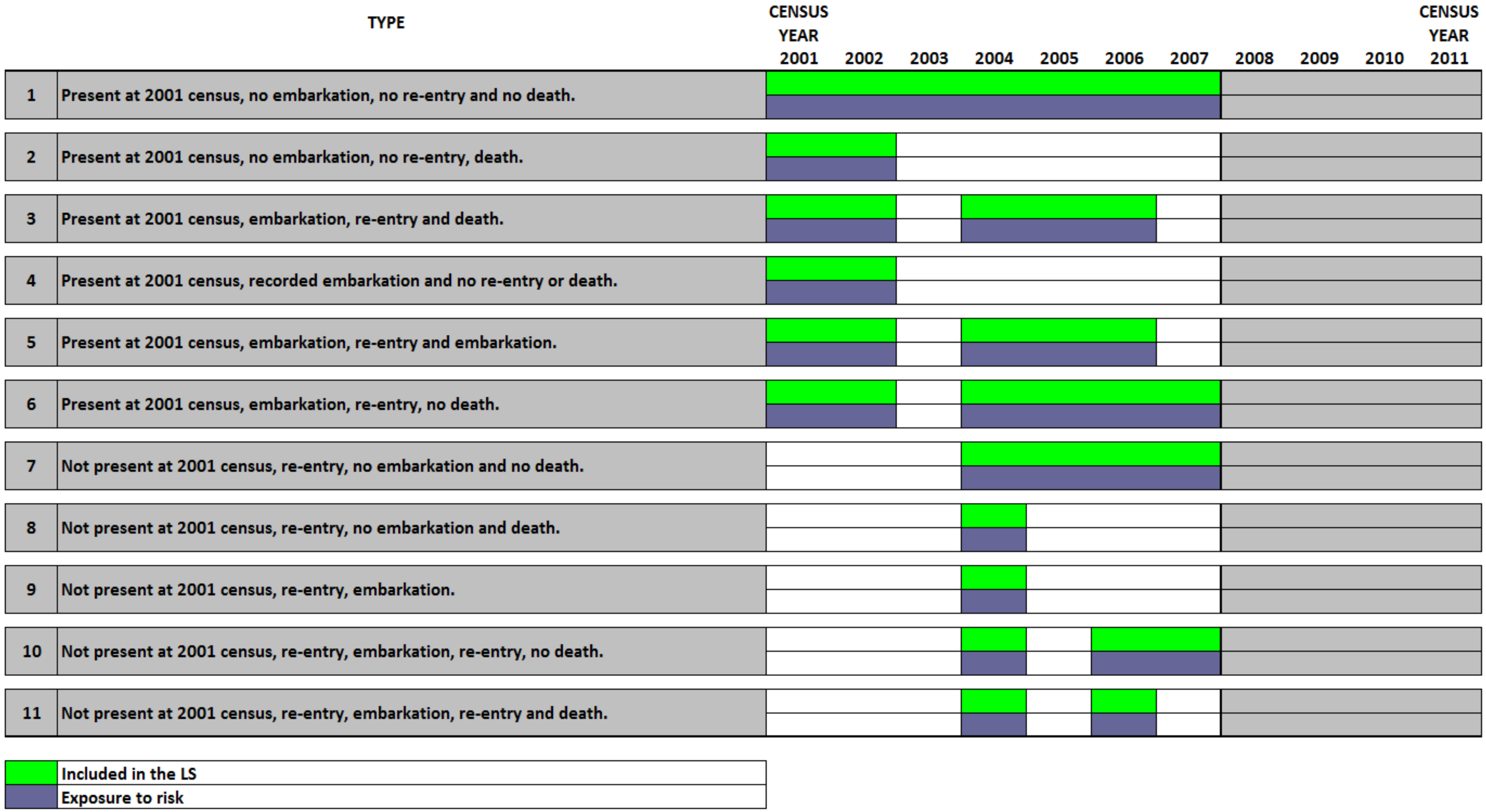
Type 9 – LS members not at the 2001 census who re-entered the study after the census, then embarked and re-entered again and died.

Type 10 – LS members not at the 2001 census who re-entered the LS, made an embarkation, re-entered and then were at the 2001 census.

Type 11 – LS members not at the 2001 census who re-entered the LS, made an embarkation, re-entered and then were recorded as dying before the 2001 census (where they were not recorded).

These types are illustrated in the following figure and numbers for each group given in Section 4.4.5.

Figure 4.10: Hypothesised consistent cases for existing members in the LS in the 2001-2007 period



Source: James Robards, July 2010.

4.4.3 Hypothesising new entrant consistent cases in the LS 2001-2007

As with the analysis above, for the 1991-2001 period, new entrants to the LS are considered separately. Six trajectories have been hypothesised and will be reported on in the next section. Included among these new entrants are LS members who come from Scotland or Northern Ireland. For all these typologies and LS members there is no census data. If they are identified after the 2011 census in the LS-Census link then their information will be included in the LS.

Type 12 – A new entrant LS member who entered the LS between 2001 and the end of 2007 and was continually resident after entry (no recorded embarkation / migration or re-entry and no record of a death).

Type 13 – LS members who entered in the 2001-2007 period, where there was no embarkation or re-entry, but where there was a death.

Type 14 – Similar to Type 2, these are LS members who immigrated to England and Wales in the 2001-2007 period, where there was an embarkation, re-entry and no death.

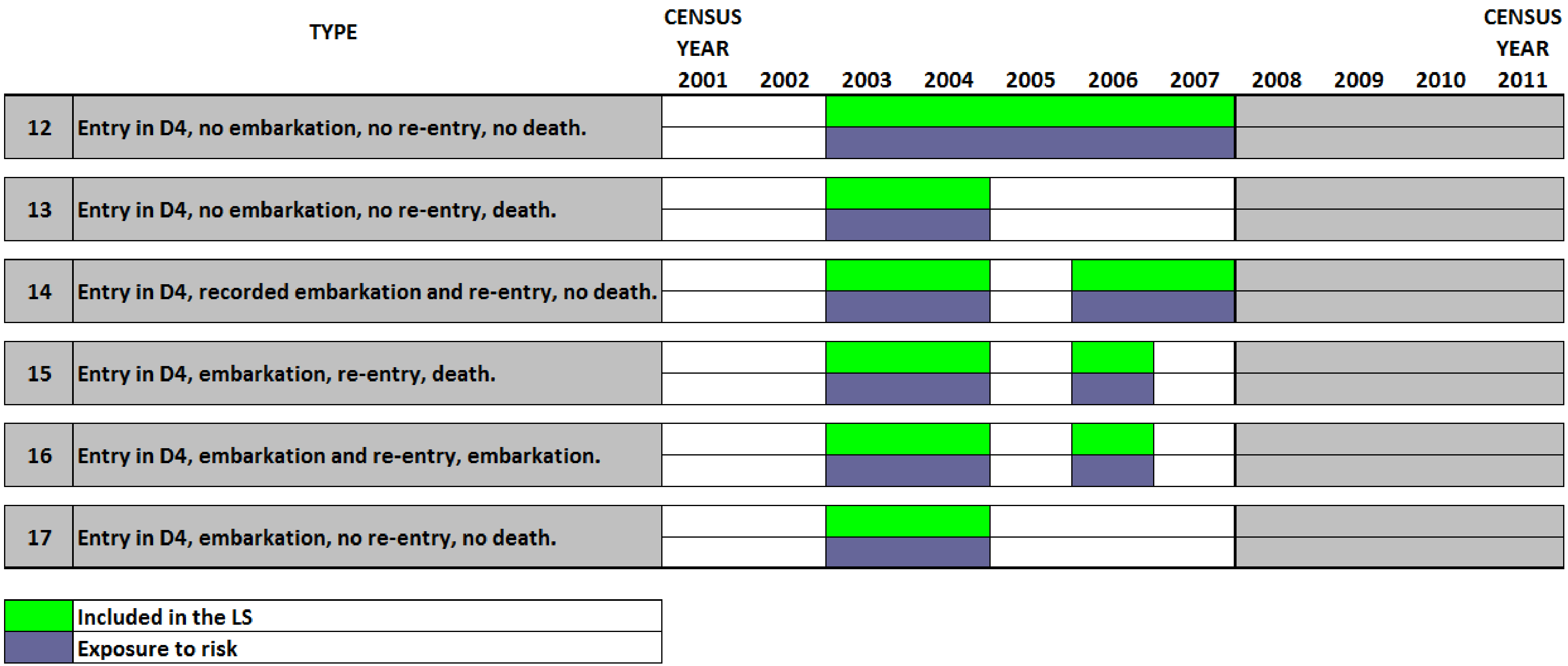
Type 15 – These are immigrants to England and Wales who entered the LS in the 2001-2007 period and where there was an embarkation, re-entry and death.

Type 16 – LS members who entered between 2001 and the end of 2007, and where there was no embarkation or re-entry, but where there was a death.

Type 17 – LS members who entered between 2001 and the end of 2007, and where there was no embarkation, but where there was a death.

These types are illustrated in the following figure, and numbers for each group given in 4.4.6.

Figure 4.11: Hypothesised consistent cases for immigrants into the LS in the 2001-2007 period



Source: James Robards, August 2011.

4.4.4 Results for consistent cases in the LS 2001-2007

Results for consistent cases in the LS since 2001 are analysed in a similar way as above for the 1991-2001 period.

4.4.5 Results for consistent existing member cases in the LS 2001-2007

Table 4.6 shows the number of LS members in the LS falling into the hypothesised trajectories for the 2001-2007 period. Again, the data is split into those cases which are new entrants into the LS since the census in April 2001 and those LS members who were are already part of the study (and may be immigrants from an earlier period). Most of the cases in the table are existing LS members in 2001 and there are high numbers of LS members in Type 1 – as a percentage of all the cases in the decade, 86% of LS members born after 1937 fall into this type. However, as was found in the analysis for the 1991-2001 period, for the fourth decade of the study it will only be at the 2011 census that LS members who have gone missing in the intercensal period can be identified.

Table 4.7 shows the percentages of existing member consistent cases in the 2001-2007 period based on being at the 2001 census. Overall, for those LS members at the 2001 census, 97.8% of the 1937-2001 birth cohorts have not shown any recorded embarkation, re-entry or death as of 2007. By birth cohort there is a slight trend where the older cohorts show higher rates of death – visible in Type 2 cohorts 1937-1956 and the slightly lower rates in the Type 1 column. Higher rates of embarkation and re-entry are shown in Type 4 for the 1962-1981 birth cohorts. This fits with the higher degree of residential mobility in these cohorts. Figure 4.12 shows the high number of consistent cases in Type 1 (non-immigrant, continuously resident) and the other categories with high enough percentages to feature in the graph. The contrast between this graph and the corresponding graphs for 1991-2001 (Figure 4.5 / Figure 4.6) is stark. Relative to the same table for the third decade (Table 4.2), there are smaller numbers in the other, not consistently resident Types. For all members the most numerous types in each part of the table are continuous residence from 2001 or entry to England and Wales.

Table 4.6: Numbers of LS members falling into consistent cases by migration status 2001-2007

Migrant status	*Existing members*								*New entrants*					*All*
BIRTH COHORT	Type 1	Type 2	Type 4	Type 6	Type 7	Type 8	Type 10	Total	Type 12	Type 13	Type 14	Type 16	Total	TOTAL
2007 - 2011	*	*	*	*	*	*	*	0	189	*	*	*	192	192
2002 - 2006	*	*	*	*	*	*	*	0	1,159	*	*	46	1,214	1,214
1997 - 2001	12,113	*	115	33	16	*	*	12,286	1,297	*	*	50	1,351	13,637
1992 - 1996	15,933	*	89	22	19	*	*	16,071	1,159	*	*	27	1,186	17,257
1987 - 1991	17,644	13	48	17	26	*	*	17,748	2,154	*	*	54	2,211	19,959
1982 - 1986	16,141	20	44	28	24	*	*	16,257	6,159	*	*	181	6,369	22,626
1977 - 1981	15,475	24	122	50	54	*	*	15,725	6,941	*	*	270	7,242	22,967
1972 - 1976	16,437	26	147	55	59	*	*	16,724	4,075	*	*	172	4,268	20,992
1967 - 1971	20,056	86	187	34	53	*	*	20,416	2,162	*	*	85	2,264	22,680
1962 - 1966	21,191	109	147	31	53	*	*	21,531	1,285	*	*	33	1,323	22,854
1957 - 1961	19,380	174	66	32	38	*	*	19,690	855	*	*	31	891	20,581
1952 - 1956	17,459	250	58	23	27	*	*	17,817	664	*	*	14	681	18,498
1947 - 1951	18,473	424	81	14	27	*	*	19,019	473	*	*	21	497	19,516
1942 - 1946	16,277	657	89	15	28	*	*	17,066	348	*	*	24	376	17,442
1937 - 1941	12,851	891	67	*	14	*	*	13,829	286	*	*	17	306	14,135
Total	219,430	2,691	1,260	360	438	*	14	224,198	29,206	24	137	1,028	30,395	254,593

**Numbers deleted to allow clearance from Micro-data Analysis and User Support (MAUS), ONS.*

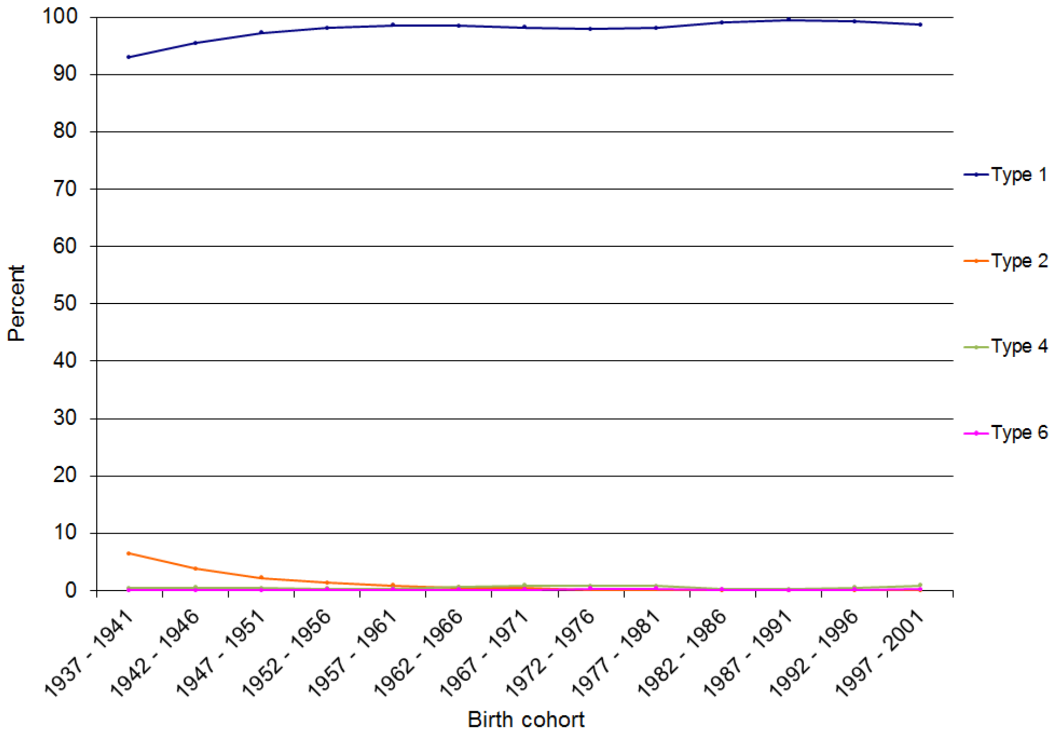
Source: ONS LS / James Robards, July 2010.

Table 4.7: Percentages of existing member consistent cases 2001-2007 who were at the 2001 census

Migrant status	*Existing members*			
BIRTH COHORT	Type 1	Type 2	Type 4	Type 6
1937 - 1941	93.0	6.4	0.5	0.0
1942 - 1946	95.5	3.9	0.5	0.1
1947 - 1951	97.2	2.2	0.4	0.1
1952 - 1956	98.1	1.4	0.3	0.1
1957 - 1961	98.5	0.9	0.3	0.2
1962 - 1966	98.5	0.5	0.7	0.1
1967 - 1971	98.2	0.4	0.9	0.2
1972 - 1976	97.8	0.2	0.9	0.3
1977 - 1981	98.1	0.2	0.8	0.3
1982 - 1986	99.1	0.1	0.3	0.2
1987 - 1991	99.5	0.1	0.3	0.1
1992 - 1996	99.2	0.0	0.6	0.1
1997 - 2001	98.6	0.1	0.9	0.3
Total	97.8	1.2	0.6	0.2

Own elaboration based on ONS LS, July 2010.

Figure 4.12: Percentage of LS members by a existing member consistent type and at the 2001 census



Own elaboration based on ONS LS, July 2010.

4.4.6 Results for consistent new entrants in the LS 2001-2007

There seems to have been a high degree of residential stability among new entrants to the LS since 2001. Type 12 is that where there was no embarkation or re-entry to the LS since the initial entry. This is the category with the largest number of people in it. Again, with 2011 census data it would be likely that some of the members in this group have 'dropped out' at some point through an unrecorded embarkation or through failing to appear on a census form. The analysis of the members who did this in the third decade give an indication of who is most likely to do this at 2011. For new entrants into the LS since 2001 there is no figure against which a benchmark can be made. With no other data on the number of LS members' resident it is not possible to look at any percentages for the LS members as a whole. Table 4.6 shows raw numbers; there are a large number of new entrants who do not show any recorded embarkations, re-entries or deaths. In the birth cohorts 1937-2011 there have been 29,206 entries since 2001 where the LS member has remained resident since entry. Other categories show smaller numbers. The second most numerous category is where there was an entry and then a death. The birth cohort with the largest number of deaths is the 1977-1981 group. These LS members would be aged between 20 and 24 years as of the 2001 census. Type 14 (recorded entry, recorded embarkation and recorded re-entry) shows small numbers – in total 137 LS members fall into this type.

4.4.7 Hypothesising existing member inconsistent cases in the LS 2001-2007

For existing members in the LS the following Types have been identified for inconsistent cases:

Type 18 – These are LS members where there has been no recorded embarkation since 2001, but where there has been a recorded re-entry. The LS member was at the 2001 census. Since the recorded re-entry there has been no recorded migration event.

Type 19 – These are LS members as in Type 1, but where there was a recorded death at some point after the recorded re-entry.

Type 20 – These are LS members where there was a recorded embarkation, no recorded re-entry and then a death.

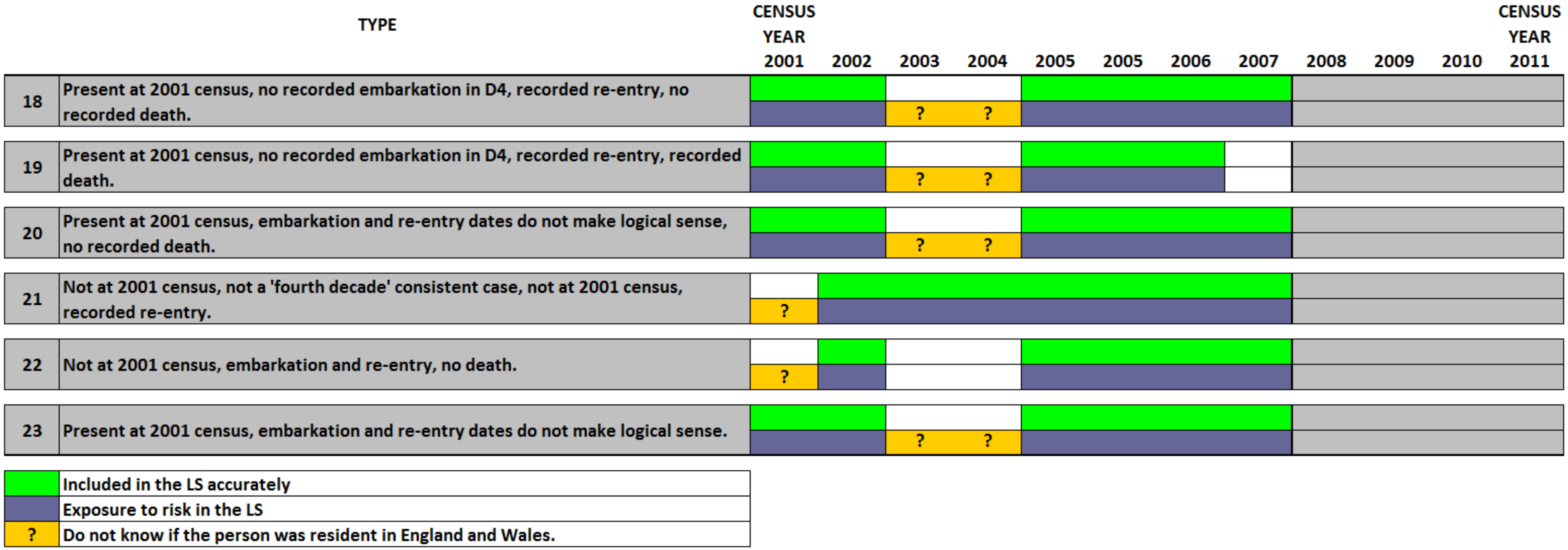
Type 21 – These are LS members who did not fall into a consistent case, were not at the 2001 census, but re-entered the LS sometime after 2001. These cases are not the same as Type 6 under the consistent cases.

Type 22 – These are LS members who were not at the 2001 census but where there was a subsequent embarkation and re-entry.

Type 23 – These are existing LS members in 2001 where the embarkation and re-entry dates do not make logical sense (i.e. on or more migration event has been missed and the degree of incompleteness is too complex to be understood and detailed).

The next section outlines the numbers of LS members falling into each of the typologies, and also the relative number in relation to the birth cohort and the population at the 2001 census.

Figure 4.13: Hypothesised inconsistent cases for existing members in the LS in the 2001-2007 period



Source: James Robards, July 2010.

4.4.8 Hypothesising new entrant inconsistent cases in the LS 2001-2007

In addition to the existing member cases, there are some forms of incompleteness for new entrants to the LS in the 2001-2007 period. In these types the incompleteness mainly relates to residence at the 2001 census and not the attrition which was identified in the 1991-2001 period. The types are shown in Figure 4.14 and explained individually:

Type 24 – These are LS members where there was a recorded entry to the LS in the period since 2001, yet where they were also recorded as being at the 2001 census. There was no other migration event recorded or death.

Type 25 – Similar to the first type, these are LS members who entered the LS since 2001 and were also recorded at the 2001 census. However, in this type there was a recorded embarkation at some point after the entry to the study.

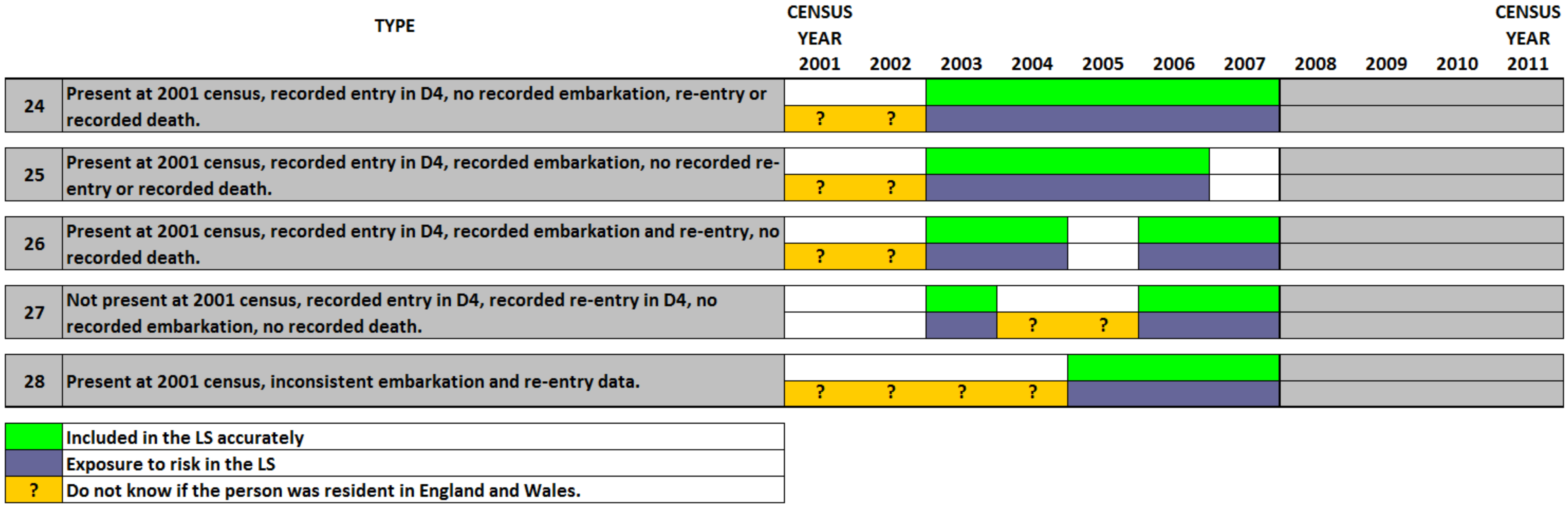
Type 26 – These are LS members who entered the LS since 2001 yet were also recorded at the 2001 census. For these LS members there was a subsequent embarkation and re-entry to the LS. The data suggests that the LS member was resident to the end of the period of observation.

Type 27 – These are LS members who re-entered in the period after the 2001 census and were not recorded at the census. They have a recorded re-entry at some point in the post-2001 census period.

Type 28 – These are LS members where there is a high degree of incompleteness or inconsistency in the embarkation and re-entry data.

Numbers are attributed to each of these types in 4.4.12.

Figure 4.14: Hypothesised inconsistent cases for new entrants in the LS in the 2001-2007 period



Source: James Robards, July 2010.

4.4.9 Results for inconsistent cases in the LS

As outlined for the consistent cases analysis, given the lack of an 'end point' to the decade, at the current time, this analysis ends in 2007. Results for existing members and new entrants are explained. Table 4.8 shows the numbers of inconsistent female cases in the LS in the 2001-2007 period by migration status since 2001.

4.4.10 Results for inconsistent existing member cases in the LS 2001-2007

Among existing member cases in the LS the most numerous category is Type 21, which is composed of LS members who did not fall into one of the consistent cases in the period since 2001 and where there was a re-entry to the study. Across age groups there is relatively little variation with a total of just 321 cases in the 1937-2006 birth cohorts. Type 18, cases where the LS member was at the 2001 census but then where there was a recorded re-entry to the study, are more unevenly spread across the birth cohorts with a concentration in the 1972-1986 cohorts and, in particular, the 1982-1986 cohort. In total there are 104 cases of this type. Apart from these two Types the other categories for existing members show very small numbers and a star has been used to denote where the number per birth cohort cannot be disclosed because of small numbers. With small numbers in all the categories where the LS member was at the 2001 census it is not possible to look at the percentages of members by residence at the 2001 census.

4.4.11 Results for inconsistent new entrant cases in the LS 2001-2007

Among the new entrants to the LS since 2001 there are also small numbers falling into an inconsistent type based on all the information since 2001. The most numerous group in the new entrant inconsistent cases is where the LS member was recorded as being resident at the 2001 census and who also entered the LS after 2001 for the first time (Type 24). Across the age groups there is some variation with higher numbers of LS members in the 1972-1981 birth cohorts. Apart from this grouping, there are 20 members where there was recorded residence at the 2001 census, a recorded entry to the LS in the post-2001 period and also an embarkation and re-entry.

Table 4.8: Numbers of LS members falling into inconsistent cases by migration status 2001-2007 period

Migrant status	*Existing members*						*New entrants*					*All*
BIRTH COHORT	Type 18	Type 20	Type 21	Type 22	Type 23	Total	Type 24	Type 25	Type 26	Type 27	Total	TOTAL
2007 - 2011	*	*	*	*	*	*	*	*	*	*	*	*
2002 - 2006	*	*	27	*	*	27	*	*	*	*	*	27
1997 - 2001	*	*	21	*	*	21	13	*	*	*	13	34
1992 - 1996	*	*	11	*	*	11	*	*	*	*	*	19
1987 - 1991	*	*	10	*	*	15	12	*	*	*	12	27
1982 - 1986	37	*	18	*	*	55	21	*	*	*	21	76
1977 - 1981	19	*	40	*	*	59	78	*	*	*	83	142
1972 - 1976	14	*	40	*	*	54	104	*	*	*	111	165
1967 - 1971	*	*	35	*	*	44	48	*	*	*	48	92
1962 - 1966	*	*	29	*	*	36	27	*	*	*	27	63
1957 - 1961	*	*	22	*	*	27	14	*	*	*	14	41
1952 - 1956	*	*	19	*	*	22	*	*	*	*	*	30
1947 - 1951	*	*	18	*	*	21	*	*	*	*	*	25
1942 - 1946	*	*	16	*	*	16	*	*	*	*	*	23
1937 - 1941	*	*	15	*	*	15	*	*	*	*	*	19
Total	104	*	321	*	*	435	348	20	*	15	386	821

**Numbers deleted to allow clearance from Micro-data Analysis and User Support (MAUS), ONS.*

Source: James Robards, July 2010.

4.4.12 Overall results for all inconsistent cases in the LS 2001-2007

The inconsistent cases presented have very small numbers. This is not surprising given that it is only when a complete decade of data is looked at that the true picture of completeness and incompleteness in the data becomes clear. Most numerous in the types outlined are cases where there was a re-entry to the LS but not of a consistent type and cases where residence at the 2001 census was recorded, yet the LS member was recorded as entering the LS for the first time in the period since 2001.

4.4.13 Conclusions for trajectories in the fourth decade of the LS

The analysis presented shows that most LS members are of a consistent type in the post-2001 time period. Table 4.6 showed that the most numerous categories of consistent cases were where there was residence at the 2001 census and no subsequent movements, or a new entry to the LS since 2001 and no movement. The birth cohort with the greatest number of new entrants' post-2001 is the 1977-1986 cohort, which is one of the key childbearing ages. Apart from deaths since the 2001 census or entry to England and Wales since 2001, there are relatively small numbers in the other types. As mentioned throughout the analysis, the lack of an 'end point' for the time frame makes the analysis of inconsistent cases more complex and somewhat tentative. However, the numbers identified are small, but useful for onward analysis. Among the new entrant cases, residence at the 2001 census seems to be the most numerous problem. This analysis and the coding of the dataset accordingly allows for the selection of LS members based on their residence and the removal of inconsistent cases if desired.

4.5 Modelling attrition from the ONS LS between the 1991 and 2001 census

LS members with certain socio-economic characteristics show a higher rate of attrition between one census and the next. Here attrition is referred to as Type 20 (1991-2001) where there was an unrecorded exit from the study between 1991 and 2001 or the LS member was not at the 2001 census or not recorded for some reason. Chapter 3 discussed the findings of ONS LS census-to-census linkage reports which are produced after each census (e.g. Blackwell et al., 2003). Through the identification of different LS member trajectories, this chapter has identified that women in their twenties are especially likely to make an unrecorded exit between 1991 and 2001 or not appear on a census form at the 2001 census, the two ways in which LS members make an unrecorded embarkation. Indeed, those LS members who were lost between 1991 and 2001 are the largest type of inconsistent case in the 1991-2001 period. Given this finding, and the past information on the characteristics most associated with attrition, it is possible to use a logistic regression model to estimate the relative weight of different socio-economic factors for these losses. The next sub-section outlines the characteristics associated with unrecorded embarkation / lost to follow-up, before section 4.5.2 shows the results from a logistic regression using some of these characteristics.

4.5.1 Hypothesising potential sources of error in the LS for fertility research

Chapter 3 focused on how the LS was initially created, how it is updated and the findings of those who have used the data before. Given the way in which the LS is constructed using the census data and the findings in reports on census to census linkage (Blackwell et al., 2003 and see also discussion in last chapter), it is possible to hypothesise the main sources of error in the LS which might lead to a less accurate representation of the population of England and Wales. For robust analysis of fertility using the LS, it is imperative that the sample exposed to risk of giving birth can be accurately identified and understood from the outset.

This sub-section is based around a detailed table of potential sources of error in the LS. Table 4.9 is a condensed version of information on attrition in the LS from readings in Chapter 3, and considers the types of individuals that are underrepresented in longitudinal samples from the LS. It covers survey design issues (problems inherent in the design of the LS and its functioning) and sample selection (the selection of a sample from the LS for fertility analysis). For attrition from the LS it is useful to have this table of characteristics derived from ONS LS reports on census-to-census linkage as a starting point. The final table shows the characteristics associated most with attrition from the LS between 1991 and 2001.

Based on Table 4.9, the following research question can be posed:

4. What are the most important socio-economic characteristic of attrition between 1991 and 2001 for female ONS LS members?

There are a total of 17,526 female LS members aged 18-85 as of 1991 who, make an unrecorded embarkation from the ONS LS at some point in the 1991-2001 period. We use the socio-economic characteristics of the LS members at 1991 to see which are statistically associated with attrition.

Table 4.9: Characteristics associated with attrition between one census and then next in the ONS LS

Group	Specific issues
Young adults	- Less likely to have events traced. - Less likely to be linked between the census dates. Unreported embarkation was a higher risk (particularly for young men). - The highest non-trace rates were for females in the 20-29 age group. - G.P. registrations in the young adult group (20s) are lower than for other age groups. - There is likely to be greater residential instability among this age group (a higher risk of temporary migration, 'gap years' and travel etc).
Women that are single (never married, divorced or living in a lone-parent household)	- Less likely to be traced or linked.
Economically inactive	- Specifically – students and the long-term unemployed. - Lower rates for tracing and linkage.
Foreign born population	- Apart from LS members from other parts of the United Kingdom (Scotland, Northern Ireland or the Isle of Man) there are lower rates of linkage and tracing for the foreign born population. - Census-to-census linkage rates for females are better than those for males.
Ethnic group	- Linkage failure far higher for LS members from ethnic minorities. - Black Africans least likely to be linked – In ONS reports links made to students coming to the UK to study (this is particularly true of London). Overall the rates for females in the Asian groups are similar to the Black groups. - Females are more likely to be linked than males.
People in social housing and private renting	- Persons living in this type of housing have lower tracing and linkage rates. - This links to the age group trend outlined above with the youngest more likely to be privately renting (lowest linkage rates).
Living in London	- There are lower rates of tracing between census dates for persons living in London. This is likely to be linked to a combination of the factors above – the population of London is younger, has a higher share of foreign born persons, a higher share of ethnic minorities and more households private renting and living in social housing. - Data on NHS registrations shows that there are 11% more people registered with a GP than estimated to be resident.

Source: James Robards, March 2010.

4.5.2 Using logistic regression to explore attrition from the LS

Although the characteristics in Table 4.9 are of some use for considering the sample for analysis, there is no information on the relative importance of background factors. Table 4.10 shows the outputs from a binary logistic regression for LS members who were at the 1991 census. The dependent variable is falling into inconsistent Type 20 which are the 17,526 LS members who were aged 18-85, resident at the 1991 census and then did not appear at the 2001 census because of an unrecorded embarkation, a cancellation on the NHSCR (these are not included in the LS because of data confidentiality restrictions) or not being on a census form at the 2001 census. All LS members at the 1991 census are included in the analysis, including those who fall into some other inconsistent type in the 1991-2001 period. Overall, 212,808 LS members aged 18-85 at the 1991 census have been included in the model.

Binary logistic regression is a suitable technique because the aim is to understand the likelihood of attrition based on a range of socio-economic variables which were recorded at the 1991 census and identified as being related to attrition. Table 4.10 also shows the pseudo R-square values (Cox & Snell R square and Nagelkerke R square) which summarise how much of the variability in the response variable is successfully explained by the model. The Nagelkerke R square (0.072) shows that 7.2% of the variability in attrition is explained by the model. Availability of other factors, more variables or characteristics associated with attrition which have not been included due to the lack of data may improve the R-square. The R-square suggests that it is possible that there are other factors which can explain variation but were not included into the model due to the lack of data. This is a limitation of the current model.

Table 4.10 shows the main results from the model and the odds ratio (Exp(B)) for each category. Only selected interactions have been included in the model. Because of the large sample being used, statistically significant interactions may be identified by chance; these may not necessarily be substantively important. Therefore, we use evidence from previous research and a priori expectations to decide which interactions to include.

In line with the expectation in Table 4.9, non UK born LS members are over three times (3.3) more likely to leave the LS in an unrecorded way. The probability of attrition declines with increasing age. The youngest age group (18-25 years in 1991) is that which is most likely to drop out of the study. The decrease in the likelihood of dropout for the older groups is in line with expectations, although the values for the 60-85 years age group could be higher given the lower LS linkage rates between census dates for the institutional population.

The results for marital status show that the widowed group is less likely to drop out than any of the other groups. Single people (the reference category) are the most likely to drop out, followed by the remarried and divorced. For this category the odds of drop-out are substantially lower for all groups compared with persons who were single in 1991.

For the socio-economic class, variable four groups have been used; these are professional and managerial, skilled manual / non-manual, partly-skilled / unskilled and missing / no-response. The reference category is those with a professional and managerial level social occupational class and relative to this, the missing / non-responses have a higher likelihood of drop out. This suggests that those who have barely completed a census form or inadequately detailed their job title at the 1991 census are most likely to drop out. Partly and unskilled occupations are not significantly different to those from professional and managerial occupations.

As expected all tenures apart from owner-occupation show a higher likelihood of attrition. In particular private renters in 1991 have an odds ratio 1.6 times that of owner-occupiers. Among the social renters the odds of drop out are lower at 1.2, while among the missing / no response group there is a higher likelihood, which would be expected given the lack of information which the LS members could provide in the 1991 census.

A measure of population density has been included in the model to try and identify if urban dwellers in the largest metropolitan areas are more likely to leave the LS. The smallest locality (of 0-199 persons per sq. KM) is the reference and relative to this, LS members residing in the largest area of in excess of 1,000 persons per sq. KM had an odds ratio showing they were 1.2 times more likely to drop out of the study.
The first set of interactions show the results for age and marital status. Middle and older aged persons who were widowed at the 1991 census show a higher likelihood of dropping out of the LS before the 2001 census, compared to younger single people. This could be related to retirement migration or movement into different forms of housing, perhaps including living with children or in an institution where they were not included on a census form or had details miss-reported leading to non-linkage. It should be noted that not all of these interaction terms are significant.

Among the interaction terms included in this model, the most interesting is probably of age and tenure. This shows that the 26-39 age group, who were private renting as of the 1991 census, were 1.3 times more likely to drop out of the LS. This is statistically significant at the 99% level. The only other interaction which shares the same level of statistical significance is that between LS members aged 60-85 in 1991 and not giving tenure on the census form. However, for this group the probability of attrition is lower than for the other age groups and tenure interactions.

To clarify the findings from these interaction terms, Figure 4.15 shows the estimated odds ratios of attrition for the housing tenure and age interaction terms in the model. These were calculated using the interaction terms and main effects (e.g. private renters in the 26-39 age group 1.65=Exp (0.271+0.455+-0.228)). The graphs make clear the age gradient in attrition and the way in which younger private and social renters (18-25 and 26-39 years at 2001) have a much higher likelihood of attrition from the LS, compared with the two older groups used in the analysis. In the graph the baseline is persons aged 18-25 who were owner-occupiers as of the 1991 census, all other covariates have been held at the reference categories.

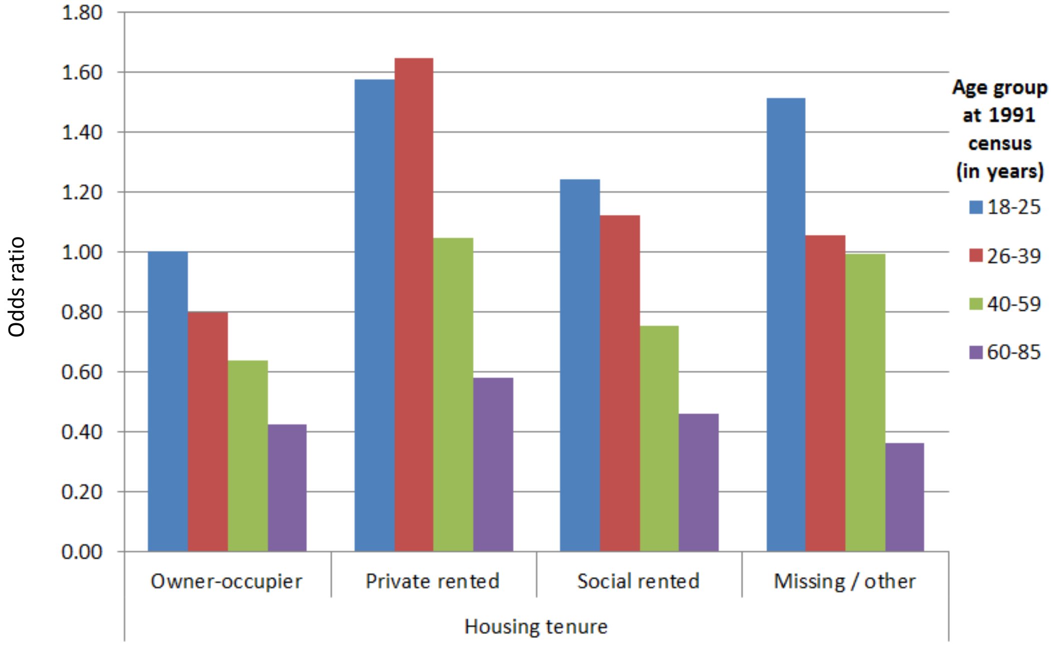
Table 4.10: Results from binary logistic regression model of the probability of attrition between 1991 and 2001 census (women, Type 20) in the ONS LS

	B	Sig.	Exp(B)
Country of birth			
UK born (reference)			
Foreign born	1.196	0.000	3.306
Age group			
18-25 (reference)			
26-39	-0.228	0.000	0.796
40-59	-0.455	0.000	0.635
60-85	-0.859	0.000	0.423
Marital status			
Single (reference)			
Married	-0.377	0.000	0.686
Remarried	-0.154	0.000	0.858
Divorced	-0.292	0.045	0.747
Widowed	-0.713	0.001	0.490
Social class			
Professional and managerial occupations (reference)			
Skilled non-manual and skilled manual occupations	-0.167	0.000	0.847
Partly-skilled and unskilled occupations	-0.017	0.691	0.983
Missing / NR	0.334	0.000	1.396
Tenure			
Owner-occupier (reference)			
Private rented	0.455	0.000	1.577
Social rented	0.215	0.000	1.240
Missing / NR	0.415	0.000	1.514
Population desnity			
0 - 199 persons per SQ. KM (reference)			
200 - 399 persons per SQ. KM	0.022	0.618	1.023
400 - 499 persons per SQ. KM	-0.134	0.039	0.875
500-999 persons per SQ. KM	-0.005	0.890	0.995
over 1000 persons per SQ. KM	0.174	0.000	1.190
Age - Marital status interaction			
Aged 60-85 years x Married	0.310	0.000	1.363
Aged 26-39 years x Divorced	0.102	0.506	1.108
Aged 40-59 years x Divorced	0.187	0.224	1.206
Aged 60-85 years x Divorced	0.301	0.086	1.351
Aged 40-59 years x Widowed	0.383	0.085	1.466
Aged 60-85 years x Widowed	0.582	0.008	1.789
Age - Social class interaction			
Aged 26-39 years x Partly-skilled and unskilled occupations	-0.143	0.010	0.866
Aged 40-59 years x Partly-skilled and unskilled occupations	-0.072	0.198	0.930
Aged 26-39 years x Missing / NR	-0.133	0.020	0.875
Aged 40-59 years x Missing / NR	-0.136	0.016	0.872
Aged 60-85 years x Missing / NR	-0.287	0.000	0.751
Age - Tenure interaction			
Aged 26-39 years x Private rented	0.271	0.000	1.311
Aged 40-59 years x Private rented	0.046	0.529	1.047
Aged 60-85 years x Private rented	-0.146	0.069	0.864
Aged 26-39 years x Social rented	0.128	0.029	1.136
Aged 40-59 years x Social rented	-0.046	0.455	0.955
Aged 60-85 years x Social rented	-0.139	0.020	0.870
Aged 26-39 years x Missing / NR	0.346	0.009	1.413
Aged 40-59 years x Missing / NR	0.032	0.834	1.032
Aged 60-85 years x Missing / NR	-0.568	0.000	0.567
Constant / intercept	-2.200	0.000	0.111

Nagelkerke R Square = .072, Cox & Snell R Square = .031, Sample = 212,808.

Own elaboration based on ONS LS, February 2012.

Figure 4.15: Estimated odds ratios of attrition by age and housing tenure from binary logistic regression model of attrition from ONS LS



Own elaboration based on ONS LS, February 2012. (Covariates held at reference category).

Foreign-born female LS members are three times more likely to leave the LS in an unrecorded way; this is a highly important consideration if we wish to use the ONS LS to estimate the fertility of recent migrants to England and Wales. This means that it would be relatively easy to have a sample of migrants which is exposed to risk of giving birth, but where we cannot be absolutely certain that the migrant is actually resident in England and Wales unless census dates are used to verify residence of LS members. Age at 1991 shows a decrease in the chance of making an unrecorded embarkation with increasing age. Women who are married, remarried, divorced or widowed all have a substantially lower odds of being lost when compared with LS members who were reportedly single at the 1991 census. In particular, those women who were widowed at the 1991 census had a lower risk of attrition 1991-2001.

Relative to the professional and managerial occupational group, the skilled manual and non-manual group and the partly-skilled and unskilled occupation group show a lower likelihood of making an unrecorded exit from the ONS LS. We do not find that economically inactive and student groups are more likely to be lost but these individuals might actually be part of the missing / no response group. For all housing tenure, owner-occupation had the lowest risk of an unrecorded embarkation. There was a sizeable difference between private rented and social rented, which is a reflection of the mobility in the private rented sector relative to the social rented.

The R-squared value suggests that there are other factors which may be included but which were not available in the data. However, despite this limitation, the results provide a better understanding of factors associated with attrition. The main aim in this case has been to explore associations with attrition.

4.6 Conclusions – selecting a sample from the LS for fertility analysis

This chapter has developed conclusions from Chapter 3 on the quality of the LS into research questions and understand the longitudinal residence of LS members and systematic bias in this. Through using residence trajectories in the 1991-2001 and 2001-2007 periods, the differential following of LS members through time has been fully explained. In England and Wales there is no other dataset that follows so many individuals over time, and thus makes the potential recording of changes among small population sub-groups possible. However, for such analysis to be undertaken, consideration of the nature of LS members continuously resident is required. Through assessing where LS members are accurately 'exposed to risk' it has been possible to code the dataset for onward analysis, while understanding systematic trends by age group.

Results for consistent cases in the LS show the number of cases for which it is possible to be completely sure where an LS member was in the third decade of the LS. The lack of information on an embarkation or other event in this decade does not mean that no such event occurred. Given the information available and the lack of a record of any such occurrence, it is only realistic to hypothesise that the LS member was resident as suggested. Of the 276,801 female LS members resident and traced at the 1991 census, it is known where a total of 90% of these members are in the third decade. This varies greatly by age group though, and the type of case that is being assessed. Among the consistent cases, overall we know that most of those LS members were continuously resident and were at the 2001 census.

However, when we analyse residence trajectories by age group it is clear that the more recent birth cohorts have fewest members that are resident continuously between (and at) the two census dates without any form of embarkation. The 1972-1981 birth cohort has a high percentage of LS members who are not continuously resident in the third decade. These members would be aged between 10 and 19 years in 1991 and 20 and 29 years in 2001. In the 1972-1976 birth cohort there is a high proportion of members which entered the study at some point after the 1991 census. Although this cohort has lower numbers of LS members continuously resident between the two census dates, it has a higher number of women entering in these years. This shows that it is necessary to consider the range of persons entering and leaving the LS. There is a high degree of flux in terms of the movements of the individuals in this age group in and out of England and Wales. Overall, the 1972-1976 birth cohort is that with the lowest percentage of consistent cases based on the trajectories devised.

When syntax was run for the inconsistent cases in the LS, many of the trajectories created had small numbers. By far the largest group in the data was for those where the LS member was at the 1991 census, not at the 2001 census, and where there was no recorded embarkation, re-entry or death. In short, the point at which the person left the study is not known. It could be that the person was resident to the census but not identified there, or they may have left at an earlier date. Many of the inconsistent cases in the LS contained small numbers, especially when disaggregated by age group.

Past reports on the tracing and linkage of LS members have said that the most problematic age group at the 2001 census was those aged 20-29 years. This corresponds with the 1972-1981 cohort, which has a high degree of incompleteness in their whereabouts for the third decade. Findings made using the hypothesised trajectories link with the reports on LS data quality and come to similar conclusions -

the birth cohort from 1972-1981 is difficult to represent accurately longitudinally with the LS. There is a need for additional research on this cohort and this will be completed in the next section, which is concerned with the number of women in the LS in each year through the 1990s and also the births to these women.

Findings for the 2001-2007 period are more tentative, with the lack of a census with which to link member trajectories. The consistent cases in this time period have very large numbers relative to the inconsistent types. However, based on the 1991-2001 findings, it is likely that a percentage of the LS members that seem to be resident at the current time might not be after the 2011 census. Through an unrecorded embarkation or failing to appear on a census form at 2011 it is likely that the final percentages of LS members continuously resident will be much lower. The most numerous non-immigrant inconsistent case at the current time is where there was a re-entry since 2001 and there is some form of incompleteness in the residence information for that member. Although there are small numbers of inconsistent histories, the coding of the dataset in this way is beneficial for post-2001 sample selection and analysis. As an indication of cases where there seems to be accurate exposure to risk, this is helpful.

Socio-economic factors which are important in accounting for attrition from the LS have been highlighted in post-census analysis following the LS-census link. Through the identification of the large number of LS members who embark in an unrecorded way between 1991 and 2001 it has been possible to understand the relative importance of different socio-economic characteristics in the attrition. The binary logistic regression model generally confirmed the hypothesised factors in Table 4.9 as leading to attrition. Key characteristics associated with attrition were being foreign born and not answering or having a missing socio-economic class response at 1991. In terms of tenure, the high attrition of private renters was evident. For use of the LS post-2001 these characteristics should be considered important when selecting a sample.

Through understanding the residence of individuals for the whole decade (1991-2001 and 2001-2007), it is possible to use the LS for years in-between with a higher degree of certainty. The residence of an LS member in a single year must be understood by examining the longer-term residence between two census dates. Having coded residence trajectories for the 1991-2001 and 2001-2007 periods, it is possible to establish the degree to which fertility rates from the LS are comparable with official statistics. The next chapter is concerned with validating the LS membership numbers for years 1991-2001 against ONS official fertility rates, official birth registration data and mid-year population estimates. Through this exercise, areas where the LS under- and over- samples will become clearer and fully understood. The residence trajectories created here will be integral to this exercise. In this chapter it has become clearer that the LS seems overall to be stronger at capturing new entrants to England and Wales than accounting for members who leave at some point.

Chapter 5

Calculating fertility rates using the Office for National Statistics Longitudinal Study

Chapter abstract

Within literature using the Office for National Statistics (ONS) Longitudinal Study (LS) there has been no recent estimation of fertility rates using the data. This chapter is concerned with identifying the sample from the LS which can be used in fertility analysis and the comparability of rates derived to official statistics. Chapter 4 identified 'consistent' and 'inconsistent' cases which are used here in the sample selection process, so that the women exposed to risk in any given year are fully understood in the context of their exposure to risk for the whole decade. The calculation of fertility rates, comparisons with ONS mid-year estimates and ONS vital statistics on births are made.

5.1 Introduction

Chapter 4 outlined the residence trajectories that an Office for National Statistics (ONS) Longitudinal Study (LS) member can take in the 1991-2001 and 2001-2007 periods. From this chapter the number of LS members taking each form of residence in the two decades of interest has been estimated. This coded the dataset so each LS member's exposure to risk in the timeframes is fully understood. The coding of the dataset in this way offers the opportunity to calculate exposure to risk for female LS members and their fertility rates in a detailed way. In this chapter this information on the longitudinal residence of the LS members is used in the calculation of fertility rates and identification of the number of women and births in the LS cross-sectionally. Past research (as explained in Chapter 3) has mainly used residence at a census as the key criteria for selecting the sample of women exposed to risk in a whole decade (or longer time frame) for analysis. The more dynamic nature of the LS (relative to other datasets), with additions to the dataset through immigration and embarkations (through recorded migrations and deaths), is not fully utilised with such a criteria. Through the use of selected trajectories from Chapter 4, fertility rates are calculated and compared to official statistics from the ONS.

In this chapter ONS mid-year estimates and ONS vital statistics are used as a benchmark for comparing the number of women and births in the LS. The next section outlines the research questions before methodology for answering these questions is described in section 5.3, before results are presented for consistent cases in section 5.4. In addition to the consistent cases this chapter also evaluates the representativeness of selected cases where there is some form of incomplete information – this work is completed in section 5.5. The LS is potentially a suitable dataset for capturing the entry of immigrants into the LS. Therefore, both consistent and inconsistent cases are used in section 5.5 to establish fertility rates, women exposed to risk and births for selected countries of birth which are comparable with the ONS 'Births and patterns of family building England and Wales' (FM1) volume.

5.2 Research questions

Residence trajectories identified in Chapter 4 and subsequent coding of the dataset allow the calculation of fertility rates based on the type of residence trajectory in the decades of interest. There has been no recent work calculating fertility rates from the ONS LS, and, as was explained in Chapter 3, there has been little work comparing the precise sample for analysis to ONS mid-year estimates and vital statistics. There is thus a gap in information about how representative the LS is for year-on-year analysis of fertility. This chapter will compare the number of LS members with the ONS mid-year estimates and ONS vital statistics figures to show where the LS is not representing either the denominator or numerator correctly, and therefore where there are representational problems.

This chapter is concerned with answering the following questions:

1. **How do age-specific fertility rates (ASFRs) from the ONS LS compare to official statistics ASFRs in the 1991-2001 period?**
   Calculating the ASFRs from the ONS LS is necessary for two reasons: firstly, this gives an overview on the suitability of the LS for fertility analysis and secondly, this tells us about the degree to which the LS reflects changing fertility rates through time.

2. **How does female membership of the LS vary year-on-year compared with ONS mid-year population estimates?**
   Of concern here is how the denominator (number of women) in the LS varies in relation to the official population statistics for England and Wales. Figures from the LS should not be expected necessarily to correspond with those from the ONS, but they will enable a comparison to be made.

3. **How do births to LS members compare with ONS vital statistics numbers?**
   Births to LS members are the numerator for calculating fertility rates. The LS offers an advantage over the calculation of fertility rates with national level statistics, as with the LS we know the population at risk, their (almost) exact age and therefore work with a sample where we know who does and who does not give birth, compared with national statistics where all births are divided by all women of a corresponding age. This is advantageous as it should provide greater stability in the denominator than is found at a national level.

4. **How well does the LS represent births to foreign born women?**
   For LS members resident at a census there is information on their country of birth from the census questions on country of birth.

There are two main reasons for establishing the accuracy of the LS's age-specific and total fertility rates – firstly to understand the degree to which the LS data is accurately reflecting births in England and Wales, and secondly to assess the dataset's suitability for the calculation of fertility behaviour among specific groups within the population.

5.3 Method

There are four main elements to the method for answering the above research questions.

5.3.1 Calculation of age-specific fertility rates and the total fertility rate

Age-specific fertility rates (ASFRs) show the intensity of fertility for different, typically five-year, age groups. To establish an ASFR, the number of women who are at a particular childbearing age in a year must be used as the denominator, and then the births to women of this age used as the numerator. The number of births in the year to women aged *x* is divided by the mid-year population of women aged *x* at their last birthday. By multiplying this figure by 1,000 the ASFR per 1,000 of the population is derived. Although this is quite straightforward with most data sources, in the case of the LS this is more complex as the composition of the LS in terms of women (and therefore births) is always changing. The filtering process used to select the sample in each year is outlined in the next sub-section.

The age-specific fertility rate is defined as:

$$fx = \frac{\textit{births in year t to women aged x last birthday at the last time of birth}}{\textit{mid-year population of women aged x last birthday}} \; x\ 1{,}000$$

(adapted from Hinde, 1998 p.100).

Total fertility rates are calculated by summing the ASFRs, dividing by 1,000 and multiplying by five (if five-year age groups have been used).

5.3.2 Selecting a sample from the LS for a numerator and denominator

As already suggested in the last sub-section, because the LS data extract provided has persons entering and leaving, it is in a constant state of flux. Chapter 3 outlined the way in which persons enter and leave the LS. The LS has over a million ever-present members (i.e. persons who have been part of the 'traced' part of the dataset at one time or another). In order to calculate ASFRs it is necessary to take into account the individual residence information for an LS member in a precise year. The analysis presented in this document has excluded individuals when they are not present in the calendar year under investigation.

As an initial step in the analysis of the LS data, and to test its suitability for analysing contemporary fertility change, age-specific fertility rates and a total fertility rate were calculated for the years 1991-2001 and 2001-2007 inclusive. To establish rates for each age group, the following filter was applied to the dataset. A sample from the LS with the following characteristics was selected:

1. Select females

- Denominator and numerator (births to female LS members).

2. Remove persons who have died

- Denominator

Exclude persons who have died by the mid-part of the year (i.e. 30 June), or in any year before that of interest.

- Numerator

Exclude persons who have died by the end of the year (i.e. 31 December), or in any year before that of interest.

3. Select persons who have been traced on the NHSCR

- Denominator and numerator

Select those persons who have ever been traced on the NHSCR and to whom a birth can be linked. If a person has not been traced on the NHSCR, event information cannot be linked to them.

4. Select persons who have entered the LS

- Denominator and numerator

Select those persons who are not immigrants (i.e. born into the study or picked up at a census) and those persons who are recorded as entering the study (i.e. through migration and registration with a GP) before the mid-year (i.e. 30 June) in the year of interest.

5. Select persons of childbearing age

- Denominator and numerator

As a general filter to reduce the size of the dataset for processing, women born after 1941 were selected. These women would be 50 in 1991. The primary objective of this filter is to remove cases from the data processing where the LS member is not in what is typically considered to be the fecund age range. For years after 1991 the appropriate denominator for the childbearing years 15-49 was selected.

6. Take into account all embarkations and re-entries to establish residence in the year

There are multiple embarkations and re-entries for some LS members. To establish who was resident in a year it is crucial to use all information for each individual. So, for each embarkation in turn:

- Denominator

Step 1- Recode embarkations (i.e. 1, 2 etc) so where there was one before the mid-year (i.e. 30 June) in the year of interest this is coded (any embarkation before 30 June in the year of observation is included in this step of the calculation).

Step 2 - Recode corresponding re-entries (i.e. 1, 2 etc) so where there was one before the mid-year (i.e. 30 June) this is coded (any re-entry before 30

June in the year of observation is included in this step of the calculation).

Step 3- Create a new variable on residence in year.

- Where there was no embarkation in step 1, then select.
- Where there was an embarkation in step 1 and a re-entry before the period of interest, then select.
- Where there was an embarkation in step 1 and no re-entry in step 2, then exclude.

Step 4- Combine all the residence variables for each event order (i.e. 1, 2, 3 and 4) and, where they all report that the person was resident for the year of interest, then select that person.

- Numerator

Step 1- Recode embarkations (i.e. 1, 2 etc), so where there was one before the end of the year (i.e. 31 December) this is coded (any embarkation before 31 December in the year of observation is included in this step of the calculation).

Step 2 - Recode corresponding re-entries (i.e. 1, 2 etc) so where there was one before the mid-year (i.e. 31 December) this is coded (any re-entry before 30 June in the year of observation is included in this step of the calculation).

Step 3- Create a new variable on residence in year.

- Where there was no embarkation in step 1, then select.
- Where there was an embarkation in step 1 and a re-entry before the period of interest, then select.
- Where there was an embarkation in step 1 and no re-entry in step 2, then exclude.

Step 4- Combine all the residence variables for each event order (i.e. 1, 2, 3 and 4) and where they all report that the person was resident for the year in use, then select.

SPSS syntax was written for each calendar year, based on the above criteria. It is important to note that this does not condition on being resident at a census. This means that both LS members who were at a census and those who enter at some point after a census are included in the analysis. Because of this there will not be the complete socio-economic background information for some LS members who may have entered since 1991 or 2001, or for some other reason have not been resident at a census. Not requiring that an LS member was at a census is likely to mean that the denominator is larger than would otherwise be the case. The criteria detailed can, however, be used across different groups in the LS and applied to the research questions identified.

5.3.3 Using inconsistent cases - imputing missing embarkations and re-entries

For the consistent cases, as detailed in Chapter 4, the above syntax will work because there is a consistency between the variables and they make logical sense. It is known for these cases in the LS that there is consistency in the information for each LS member. While it is not possible to know if there are some unrecorded events which have not been picked up, on the basis of the information in the LS, and therefore to the best of our knowledge, these are consistent cases. However, among the inconsistent cases identified in Chapter 4 there are missing embarkation or re-entry data of some form. Where there is some form of missing embarkation or re-entry data it is possible to impute a year of re-entry and therefore allocate exposure to risk to the inconsistent cases.

There are many options for deciding on the exposure which should be attributed to such LS members, and it is possible to investigate this in some detail. However, the simplest starting point is to allow the members who have some form of incompleteness half the exposure of the complete decade. The use of event information (e.g. dates of cancer registrations or births) to impute a date of embarkation or re-entry might lead to bias in the exposure to risk in the dataset and in subsequent calculations. Because of the nature of incompleteness for some of the typologies which have been devised, and the high level of inconsistency in their event information (i.e. embarkations after re-entries), some cases will be dropped from analysis under the consistent typologies group. Tables 5.1-5.3 show 'problem cases' in the dataset, the number of cases and the solution adopted to use these cases in analysis.

Table 5.1: Imputation of missing embarkation – inconsistent cases

Inconsistent type	***Description***	***Number of cases- birth cohorts (1937-2011)***	***Imputation solution***
Type 20	The person was at the 1991 census, they were not at the 2001 census and there is no record of an embarkation, re-entry or death.	21,382	Embarkation = 1996
Type 23	The person enters the LS and then goes 'missing' before 2001, with no record of a death or embarkation.	14,110	Embarkation = 1996
Type 25	These are LS members where the entry to the LS is recorded, and where there is an embarkation and re-entry. At some later point the LS member leaves as there is no recorded death or embarkation, or the LS member is not recorded at the 2001 census. There was no recorded embarkation after the re-entry to the LS	87	Embarkation = 1996
Total		**35,579**	

Own elaboration based on ONS LS, August 2010.

Table 5.2: Imputation of missing re-entry – inconsistent cases

Inconsistent type	*Description*	*Number of cases-birth cohorts (1937-2011)*	*Imputation solution*
Type 18	An LS member who was present in 1991 and who had a recorded embarkation in the third decade of the LS, but no recorded re-entry and then the person was identified at the 2001 census.	26	Re-entry/ entry = 1996
Type 21	These are LS members who were not resident in 1991 but who may have been resident in the LS at some point in the past. They 'appeared' at the 2001 census and it is not known when they may have entered England and Wales before this date (unless they were an LS member in the pre-1991 period).	18,362	Re-entry/ entry = 1996
Type 26	These are LS members who make a recorded embarkation but are recorded at the 2001 census. The date of re-entry to England and Wales is not known.	*	Re-entry/ entry = 1996
Total		**50,033**	

Own elaboration based on ONS LS, August 2010. Note - * denotes values deleted to meet ONS disclosure controls / no cases.

Table 5.3: Cases to be dropped from analysis – inconsistent cases

Inconsistent type	*Description*	*Number of cases- birth cohorts (1937-2011)*	*Imputation solution*
Type 22	This is a case where the LS member was at the 1991 census, but where there is an inconsistency in the recorded embarkation and re-entry data. The dates are incorrectly recorded, or there is an inconsistency where the date of re-entry was before the embarkation.	227	Drop from analysis.
Type 19	These are LS members who made an unrecorded embarkation, but a recorded re-entry.	*	Drop from analysis.
Types 24	These are LS members where there is a recorded death and yet the LS member is found to be at the 2001 census.	*	Drop from analysis.
Types 27	The LS member was recorded as entering, and then there is an inconsistency in the embarkation and re-entry dates within the third decade. The LS member is at the 2001 census.	*	Drop from analysis.
Types 28	LS members who are new entrants where there was an entry date, no recorded embarkation and then a re-entry date. This means that they were 'missing' for an unknown period from the date of entry, through to the date of re-entry.	*	Drop from analysis.
Types 29	LS members who were recorded as entering the study in the third decade, yet were also recorded as being at the 1991 census.	*	Drop from analysis.
Types 30	These are persons who are resident at one of the census dates in the third decade but where the embarkation and re-entry dates do not make logical sense, in that the embarkation is recorded as being after the re-entry.	*	Drop from analysis.
Total:		**227**	

Own elaboration based on ONS LS, August 2010. Note - * denotes values deleted to meet ONS disclosure controls / no cases.

For these imputed years of entry and embarkation to be used in the calculation of a denominator, numerator and fertility rate, the syntax as outlined in section 5.3.2 has been adjusted so as to take into account the information in the above tables. Key is 'step 6' where the embarkation and re-entry information is used. This has been adjusted so as to use the embarkation and re-entry data in the LS and also the imputed final embarkation and re-entry.

5.3.4 Comparing the rate, numerator and denominator from the LS to ONS official and vital statistics

The ASFRs, women and births derived from the ONS LS can be compared to the corresponding official statistics figures to understand the degree to which the ONS LS is accurately representing certain age groups. By making the comparisons it is possible to understand where there is an over- and under-sampling in the LS relative to the official statistics. In particular, deviations in the ASFRs can be understood through using the denominator and numerator comparisons. It will then become apparent whether there is a denominator or numerator bias in the LS and whether this is true across the dataset or only for some age groups. Data for the comparisons is readily available in the ONS historical births series (FM1) and from ONS mid-year estimates.

The key measures used in the comparison are:

1. ***'Sampling fraction of official statistics women by the LS'***

   This is the percentage of the England and Wales estimated mid-year population (30 June) of women represented by the LS (conventionally termed the 'sampling fraction').

- Calculated by:
    1. Dividing the number of LS women age x in the mid-year by the corresponding figure for England and Wales and multiplying by 100.

$$\textbf{\textit{Sampling fraction of mid-year estimate women by the LS}} = \frac{\textit{LS women at mid-year t aged x last birthday}}{\textit{ONS estimate of women at mid-year t aged x last birthday}} \times 100$$

2. ***'Representation of LS women based on official statistics'***

   This is the number of women in the LS divided by the expected women in the LS (conventionally called the 'linkage rate').

- Calculated by:
    1. Establishing the 'expected' number of women in the LS by dividing 365.25 (days of the year inclusive of 0.25 for leap years) into 4 (the

number of birth dates used in the LS) and then multiplying this by the England and Wales women figure.

2. Dividing the number of LS women by the 'expected' number of women in the LS to give the final 'representation of LS women based on official statistics' figure.

$$\textbf{\textit{Expected women in the LS}} = \frac{4}{365.25} \times \textit{England and Wales women}$$

$$\textbf{\textit{Representation of LS women based on official statistics}} = \frac{\textit{LS women}}{\textit{Expected women in the LS}}$$

3. *'Sampling fraction of official statistics births by the LS'*

This is the percentage of England and Wales official births represented by the LS (conventionally termed the 'sampling fraction').

- Calculated by:
    1. Dividing the number of LS births by the corresponding figure for England and Wales and multiplying by 100.

$$\textbf{\textit{Sampling fraction of official statistics births by the LS}} = \frac{\textit{LS births in year t to women aged x last birthday at the last time of birth}}{\textit{ONS FM1 births in year t to women aged x last birthday at the last time of birth}} \times 100$$

4. *'Representation of LS births based on official statistics'*

This is the number of births in the LS divided by the expected births in the LS (conventionally called the 'linkage rate').

- Calculated by:
    1. Establishing the 'expected' number of births in the LS by dividing 365.25 (days of the year inclusive of 0.25 for leap years) into 4 (the

number of birth dates used in the LS) and then multiplying this by the England and Wales births figure.

2. Dividing the number of LS births by the 'expected' number of expected births in the LS to give the final 'representation of LS births based on official statistics' figure.

$$\textit{Expected births in the LS} = \frac{4}{365.25} \times \textit{England and Wales births}$$

$$\textit{Representation of LS births based on official statistics} = \frac{\textit{LS births}}{\textit{Expected births in the LS}}$$

5.4 Results - Consistent cases from the LS: fertility rates, women and births

The first step in this analysis is to understand the overall fertility rates, numerators and denominators for individual calendar years in 1991-2001. Using only consistent cases means there is a high degree of certainty on the residential stability and the contribution to the denominator and numerator of these LS members through time is assured. This section outlines annual ASFRs and TFRs for all consistent cases and cases in the dataset where there was continuous residence in the 1991-2001 period – 'Type 1 consistent cases'. The same analysis has not been run for LS members post 2001, as there is a lack of certainty over how many LS members will be consistent when 2011 census data is linked in. It is only then that unrecorded embarkations and loss to follow-up can be identified.

5.4.1 Fertility rates

Tables 5.4-5.15 show the number of women, births and age-specific fertility rates for all consistent cases in the LS and all the Type 1 consistent cases (i.e. those persons continuously resident between the 1991 and 2001 census dates). The last column of each table shows the data from the FM1 'Births Statistics' volume from the ONS, which is the official rate for England and Wales.

All the tables show that rates from all the consistent cases and those from the Type 1 cases (continually resident consistent cases) are similar. There are deviations between the two types. In the childbearing years, Type 1 cases generally have lower rates and this leads to a slightly lower overall TFR for the values from the Type 1 cases. It is only in 1993, where the TFR for the Type 1 cases is greater than the all consistent cases. For all the consistent cases the TFRs are highly comparable and generally within 5% of the official figures. With the official ASFRs there are slightly lower rates across the all age groups. Sections 5.4.2 and 5.4.3 make comparisons to the vital statistics and mid-year estimates data to highlight the sources of this deviation. Figure 5.1 shows the average figures for the decade against official ASFRs for England and Wales. The difference in the TFRs for the Type 1 consistent cases and the consistent cases increases as there is a cumulative effect of summing the slightly different ASFRs in each of the age groups. This leads to the TFR which is different from the equivalent FM1 figures.

Table 5.4: Consistent cases - women, births and ASFRs for 1991

AGE	ALL CONSISTENT CASES			ALL TYPE 1 CONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	13,803	428	31.0	13,609	421	30.9	33.0
20-24	17,055	1,448	84.9	16,667	1,417	85.0	89.3
25-29	18,794	2,205	117.3	18,384	2,164	117.7	119.4
30-34	17,960	1,501	83.6	17,583	1,471	83.7	86.7
35-39	16,528	473	28.6	16,151	462	28.6	32.1
40-44	17,968	71	4.0	17,498	69	3.9	5.1
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.75			1.75	1.82

*Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.5: Consistent cases - women, births and ASFRs for 1992

AGE	ALL CONSISTENT CASES			ALL TYPE 1 CONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	13,101	390	29.8	12,896	376	29.2	31.7
20-24	16,504	1,430	86.6	16,009	1,386	86.6	86.2
25-29	18,770	2,115	112.7	18,248	2,057	112.7	117.5
30-34	18,385	1,566	85.2	17,955	1,527	85.0	87.3
35-39	16,803	519	30.9	16,404	497	30.3	33.4
40-44	17,200	96	5.6	16,751	92	5.5	5.5
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.75			1.75	1.80

*Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.6: Consistent cases - women, births and ASFRs for 1993

AGE	ALL CONSISTENT CASES			ALL TYPE 1 CONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	12,857	344	26.8	12,627	344	27.2	30.9
20-24	16,114	1,194	74.1	15,533	1,185	76.3	82.6
25-29	18,529	2,122	114.5	17,935	2,104	117.3	114.4
30-34	18,662	1,494	80.1	18,182	1,475	81.1	87.3
35-39	17,199	578	33.6	16,797	569	33.9	34.1
40-44	16,794	106	6.3	16,366	104	6.4	5.9
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.68			1.71	1.77

*Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.7: Consistent cases - women, births and ASFRs for 1994

AGE	ALL CONSISTENT CASES			ALL TYPE 1 CONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	12,924	357	27.6	12,651	343	27.1	28.9
20-24	15,644	1,131	72.3	15,009	1,069	71.2	79.1
25-29	18,426	1,996	108.3	17,708	1,938	109.4	112.4
30-34	18,973	1,528	80.5	18,460	1,489	80.7	89.3
35-39	17,513	569	32.5	17,068	555	32.5	35.8
40-44	16,501	118	7.2	16,115	116	7.2	6.1
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.64			1.64	1.75

*Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.8: Consistent cases - women, births and ASFRs for 1995

AGE	ALL CONSISTENT CASES			ALL TYPE 1 CONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	13,308	336	25.2	13,048	321	24.6	28.5
20-24	15,337	1,087	70.9	14,610	1,005	68.8	76.4
25-29	17,861	1,865	104.4	17,034	1,766	103.7	108.7
30-34	19,221	1,631	84.9	18,598	1,563	84.0	88.2
35-39	17,829	596	33.4	17,358	575	33.1	36.4
40-44	16,401	113	6.9	16,016	102	6.4	6.5
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.63			1.60	1.72

*Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.9: Consistent cases - women, births and ASFRs for 1996

AGE	ALL CONSISTENT CASES			ALL TYPE 1 CONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	13,714	374	27.3	13,411	364	27.1	29.7
20-24	14,384	1,153	80.2	13,618	1,079	79.2	77.0
25-29	17,629	1,830	103.8	16,680	1,732	103.8	106.9
30-34	19,102	1,699	88.9	18,391	1,628	88.5	89.7
35-39	18,085	650	35.9	17,587	619	35.2	37.5
40-44	16,555	117	7.1	16,159	112	6.9	6.9
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.72			1.70	1.74

*Own elaboration based on: ONS LS, FM1 no.30, 2001, September 2010. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.10: Consistent cases - women, births and ASFRs for 1997

AGE	ALL CONSISTENT CASES			ALL TYPE 1 CONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	13,968	381	27.3	13,629	368	27.0	30.3
20-24	13,701	998	72.8	12,903	901	69.8	75.9
25-29	17,130	1,798	105.0	16,021	1,659	103.6	104.5
30-34	19,074	1,596	83.7	18,255	1,511	82.8	89.8
35-39	18,508	698	37.7	17,958	666	37.1	39.3
40-44	16,833	121	7.2	16,409	116	7.1	7.3
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.67			1.64	1.73

*Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.11: Consistent cases - women, births and ASFRs for 1998

AGE	ALL CONSISTENT CASES			ALL TYPE 1 CONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	13,913	437	31.4	13,528	415	30.7	31.2
20-24	13,503	995	73.7	12,628	896	71.0	74.9
25-29	16,804	1,621	96.5	15,540	1,505	96.8	101.6
30-34	18,897	1,614	85.4	17,939	1,515	84.5	90.7
35-39	18,808	695	37.0	18,186	649	35.7	40.4
40-44	17,230	124	7.2	16,798	114	6.8	7.5
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.66			1.63	1.73

*Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.12: Consistent cases - women, births and ASFRs for 1999

AGE	ALL CONSISTENT CASES			ALL TYPE 1 CONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	14,158	465	32.8	13,748	436	31.7	31.1
20-24	13,606	900	66.1	12,651	791	62.5	73.1
25-29	16,429	1,643	100.0	15,010	1,485	98.9	98.4
30-34	18,841	1,608	85.3	17,710	1,491	84.2	89.7
35-39	19,166	727	37.9	18,461	688	37.3	40.6
40-44	17,556	104	5.9	17,068	94	5.5	7.7
45-49	*	*	<0.5	*	*	<0.5	0.4
TFR			1.64			1.60	1.70

*Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.13: Consistent cases - women, births and ASFRs for 2000

AGE	ALL CONSISTENT CASES			ALL TYPE 1 CONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	14,490	347	23.9	14,020	333	23.8	29.5
20-24	14,042	887	63.2	13,049	752	57.6	70.2
25-29	16,206	1,466	90.5	14,611	1,272	87.1	94.5
30-34	18,298	1,577	86.2	17,035	1,440	84.5	88.1
35-39	19,469	734	37.7	18,599	678	36.5	41.4
40-44	17,893	114	6.4	17,358	106	6.1	8.0
45-49	*	*	<0.5	*	*	<0.5	0.4
TFR			1.54			1.48	1.66

*Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.14: Consistent cases - women, births and ASFRs for 2001

AGE	ALL CONSISTENT CASES			ALL TYPE 1 CONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	15,002	375	25.0	16,379	350	24.2	28.0
20-24	14,593	916	62.8	16,856	794	59.2	69.3
25-29	15,360	1,492	97.1	17,454	1,292	94.9	91.9
30-34	18,145	1,593	87.8	19,622	1,421	85.2	88.2
35-39	19,396	746	38.5	20,640	682	37.1	41.6
40-44	18,173	132	7.3	19,379	121	6.9	8.4
45-49	*	*	<0.5	*	*	<0.5	0.5
TFR			1.59			1.54	1.64

*Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.15: Consistent cases - women, births and ASFRs, Average for 1991-2001

AGE	ALL CONSISTENT CASES			ALL TYPE 1 CONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	13,749	385	28.0	13,595	370	27.6	30.3
20-24	14,953	1,104	73.4	14,503	1,025	71.6	77.6
25-29	17,449	1,832	104.6	16,784	1,725	104.2	106.4
30-34	18,687	1,582	84.7	18,157	1,503	84.0	88.6
35-39	18,119	635	34.9	17,746	604	34.3	37.5
40-44	17,191	111	6.4	16,902	104	6.2	6.8
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.66			1.64	1.74

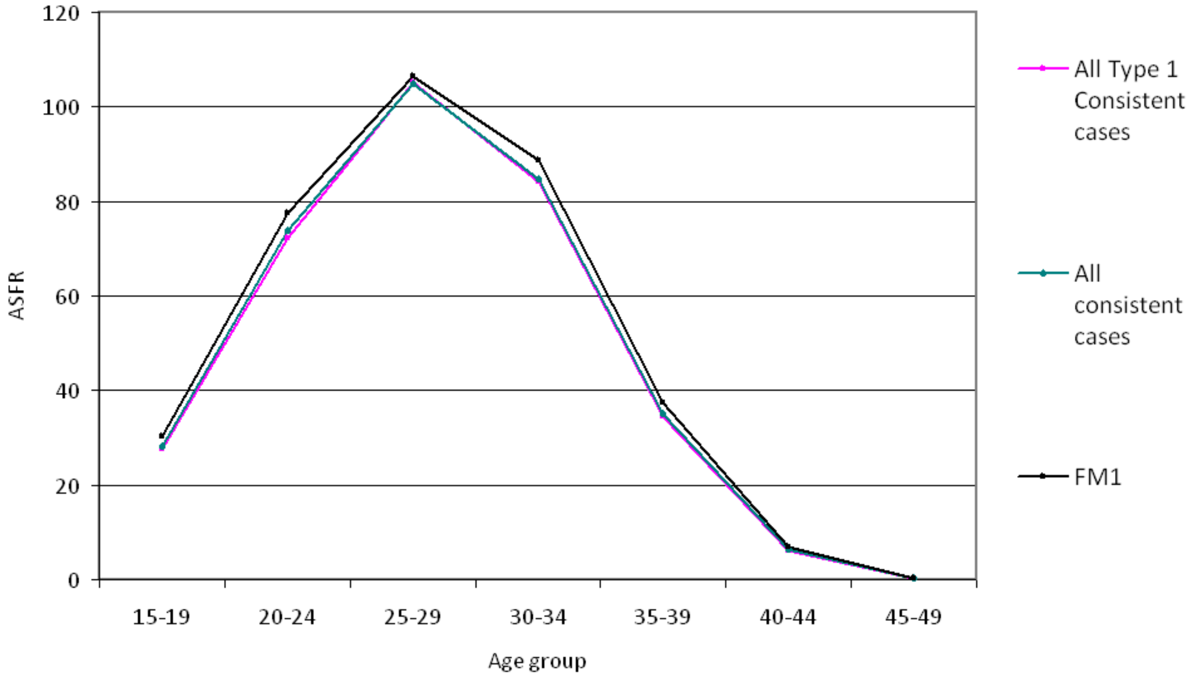
*Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010. Note - * denotes values deleted to meet ONS disclosure controls.*

Comparing LS fertility rates with Official Statistics rates

Although much can be learned from the tables above, Figure 5.1 (see appendix for single years) shows the data in graphs. This highlights the difference in the rates between all the consistent cases, consistent cases of Type 1 and the official rates. In almost all the appended graphs there is a very high degree of comparability between the overall consistent cases and the Type 1 consistent cases. Throughout the 1991-1995 period the rates from the LS are very slightly lower than the official statistics comparator. In 1996 only the 20-24 age group from the LS has a higher rate than the official figures. In 1999 there is a slightly higher rate for women in the 25-29 age group from the LS. Apart from this, the rates in this year are notably lower for the 20-24 and 30-34 age group official figures. In 2000 and 2001 there is more of a difference between the consistent cases and consistent Type 1 cases than in other years in the decade. Figure 5.1 shows the 1991-2001 ASFR average, this shows the high degree of comparability through the decade.

In later years of the 1991-2001 period the age groups 15-19 and 20-24 (as identified in the tables) show a larger gap between the official rates and the LS consistent cases rates than for other age groups. In some years (notably 1999 and 2000) the rates for the 20-24 age group are some distance from the official rate. However, overall there is a very high degree of comparability between the all consistent cases category, the Type 1 cases and the official rate. This is particularly true among the older age groups (30-34 and older) where in all the years there is a high degree of comparability. It is only in a few years (e.g. 1994 and 2000) where there is a deviation between the groups. The potential impact on the fit of the fertility rates by including selected inconsistent cases in the denominator and numerator will be interesting, given the findings here. It is possible that there will be an increase in the denominator (number of women) which may erode the rates and make the combined consistent and selected inconsistent cases rates too low in comparison with the ONS ASFRs. The denominator would be easy to get wrong through use of selected inconsistent cases and there would also be an impact on the numerator for analysis. However, the inclusion of some inconsistent cases may make the number of women in the 15-19 and 20-24 groups more representative if there is a denominator problem, shown in the next section looking at the comparability of the consistent cases with the ONS mid-year estimates.

Figure 5.1: Consistent cases - age-specific fertility rates, Average values for 1991-2001



Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010.

5.4.2 Comparing the denominator (women) with ONS mid-year estimates

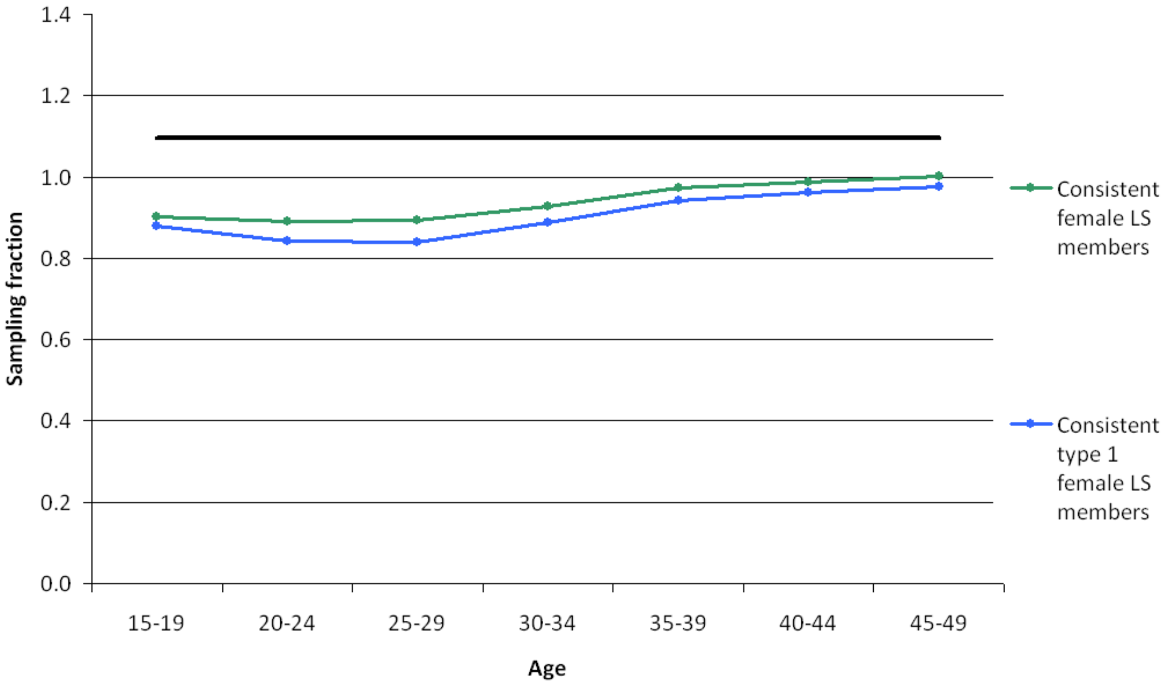
It is possible that there could be offsetting errors and therefore we look at the numerators and denominators to identify the effect of any offsetting error.

Sampling fraction of official statistics women by the LS

Through calculating the number of women in the LS at the mid point of each year, the 'sampling fraction of official statistics women by the LS' and the 'representation of LS women based on official statistics' can be established. These show how the LS compares to the 1.09% target based on the trajectories created and used in this analysis. When the same figures for births in the LS in the same time period are used, the rates show where the deviations in the ASFRs arise from the denominator or numerator. In section 5.4.1 it was hypothesised that there could be a lower denominator in the 15-19 and 20-24 year age groups than would be desired, and that it is this that might be lead to the deviation among the consistent cases from the official statistics for these age groups. As was explained in the methodology section, it is possible to compare the number of women in the LS selected in the consistent and consistent Type 1 cases to the ONS official mid-year estimates. Comparisons by single years of age give rates which fluctuate greatly, and so five-year bands are used. Figure 5.2 shows the average sampling fraction of the LS of official statistics for women in the LS for all years between 1991 and 2001. The thick black line just above 1 on the scale shows the target sampling for the LS (1.09%).

The figure shows there is a lower rate of sampling among the 15-29 age ranges for both all consistent and consistent Type 1 cases. This shows that, as hypothesised at the end of the last section, there is a problem with the denominator for consistent cases in the LS in these age groups. From 1996 through to 2001 (see appendix) there is a widening gap between all consistent cases and the Type 1 consistent cases. The gap may be arising from the new entrants to the LS in the 1990s adding to the number of LS members in some of the key childbearing age groups. For all consistent cases there is a decline from the 1.09% target from the earlier part of the decade; this is most pronounced for the 25-29 year age group. Over the decade there seems to be a cohort trend, as the low rates for the early part of the decade for the 15-19 and 20-24 year age groups filter through to the 30-34 year age group at the end of the decade. For the whole decade the 40-44 and 45-49 year age groups are those where there is a sampling fraction closer to the 1.09% target. These findings match those in the last chapter – the older age groups are better represented through time.

Figure 5.2: Consistent cases - sampling fraction of official statistics women by the LS, Average 1991-2001



Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 Local Authority Population Studies: 07/10/04, ONS Mid-year estimates, Mid-2001 (Revised) Local Authority Population Studies: 09/09/04, Accessed June 2010.

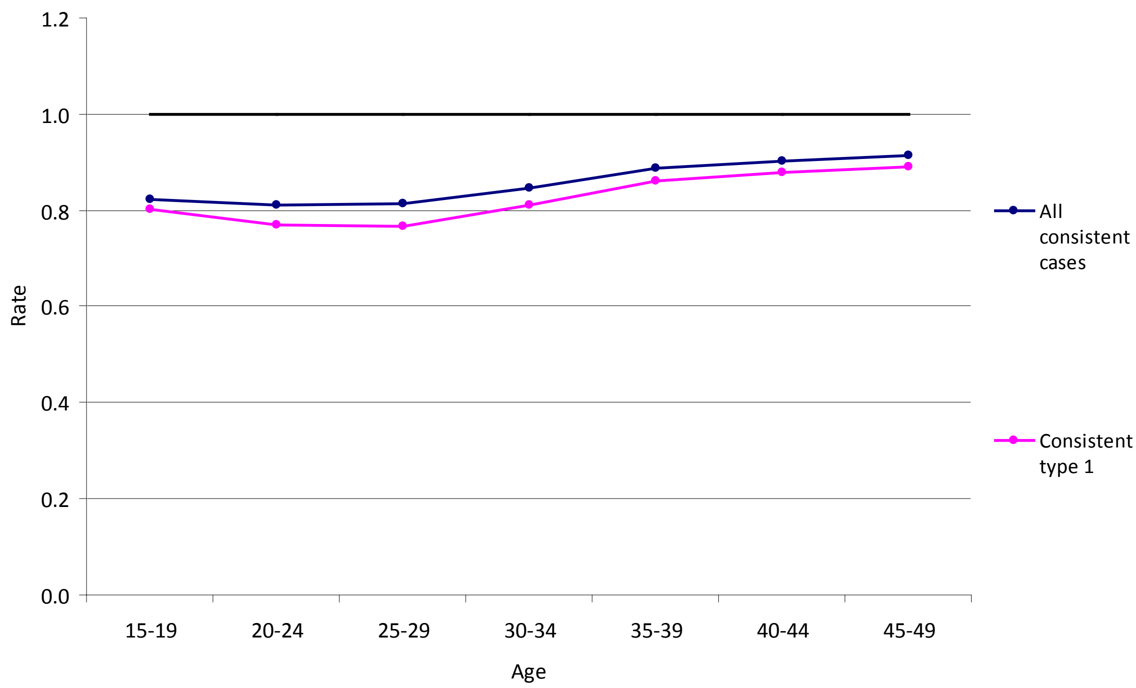
Representation of LS women based on official statistics

Figure 5.3 shows the representation of LS women based on official statistics – these values use the expected sample size of the LS, based on the corresponding population of England and Wales. The difference is that the comparison being made is to an 'expected' sample size rather than the actual number of cases which have been estimated (in the case of the 'sampling fraction of official statistics women by the LS').

On this measure all age groups throughout the decade are lower than the 1% sampling target. The profiles for each of the years are generally similar to those in the last sub-section, with increasing representation with age. Also in line with the last sub-section, the all consistent cases give a better representation with values closer to the 1% target. In the later part of the decade there is a widening gap between the women in the LS and the target value – this is in line with the sampling rates above. Generally, compared with the last section, there is a smaller variation between the all consistent and the consistent Type 1 cases than there was in the last section. This might be because a more conservative sample is expected than compared with the actual mid-year number of women estimated by the ONS.

The results for the women in the LS in this section have been placed before the analysis of the comparison of the births, as the sample of women who are exposed to risk will impact on the births in the LS and therefore the comparison of these to the ONS vital statistics official rates.

Figure 5.3: Consistent cases - representation of LS women based on official statistics, Average 1991-2001



Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 Local Authority Population Studies: 07/10/04, ONS Mid-year estimates, Mid-2001 (Revised) Local Authority Population Studies: 09/09/04, Accessed June 2010.

5.4.3 Comparing the numerator (births) with ONS vital statistics

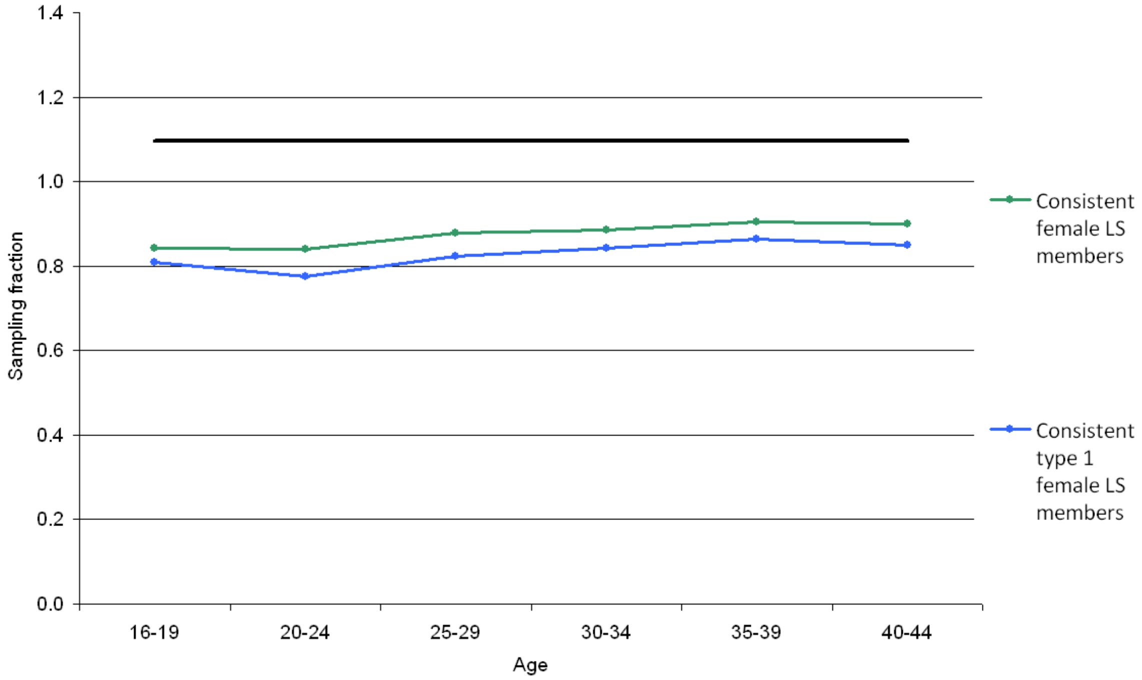
As with the last section, it is also possible to compare the number of births reported to women in the LS to the official statistics figures. Due to small numbers in some of the cells from the analysis for this section and the ONS restrictions on the reporting of small numbers (see in opening section to this document), it is not possible to run this analysis for single years of age. Five-year age groups are used in the analysis here and Figure 5.4 shows the sampling fractions of official statistics births by the LS on average for 1991 – 2001.

Sampling fraction of official statistics births by the LS

Compared with the last section on the denominator in the LS, the sampling fractions for births are more erratic, with larger differences between age groups and a sampling rate which is generally lower than for the women in the LS (in the last section). Births to older women in the data are generally best represented; however, there are variations from year to year. Figure 5.4 contains the average for the 1991-2001 period and shows there is very little variation from age group to age group in the decade as a whole. In almost all years (see appendix) the youngest age group (16-19 years) shows the lowest sampling fraction – around 0.8. The oldest and youngest age groups are vulnerable to erratic trends by virtue of the relatively small numbers of births compared with the key childbearing ages. However, the 20-24 year age group also has a similar rate in most cases. Given some of the findings in the last chapter regarding longitudinal follow up of younger members of the LS, the lower figures here should not be a surprise. In the mid to later part of the decade there is an increased difference between the all consistent cases and the consistent Type 1 cases. This matches the findings on the number of women in the LS in the last sub-section.

For 1999, 2000 and 2001 the 35-39 and 40+ year age groups decline, from around 1 - where they are for much of the decade – to a rate comparable with the 16-19 year age group. Interestingly, in the 1999-2001 period there is an increasing gap which opens up between the all consistent and Type 1 consistent cases in the dataset. Throughout the decade, and particularly towards the end of the decade, the all consistent cases category is better at sampling births. Later in the decade the deviation between the two lines may be related to the all consistent cases sampling births to new entrants into the study, whereas the consistent Type 1 cases are limited to the sample at the 1991 census. This is because the consistent Type 1 cases are those persons who are resident between the 1991 census and the 2001 census with no recorded embarkation. The same group at 1991 is therefore also at 2001 and there are no additions (through migration) or exits through deaths or embarkations.

Figure 5.4: Consistent cases - sampling fraction of official statistics births by the LS, Average 1991-2001

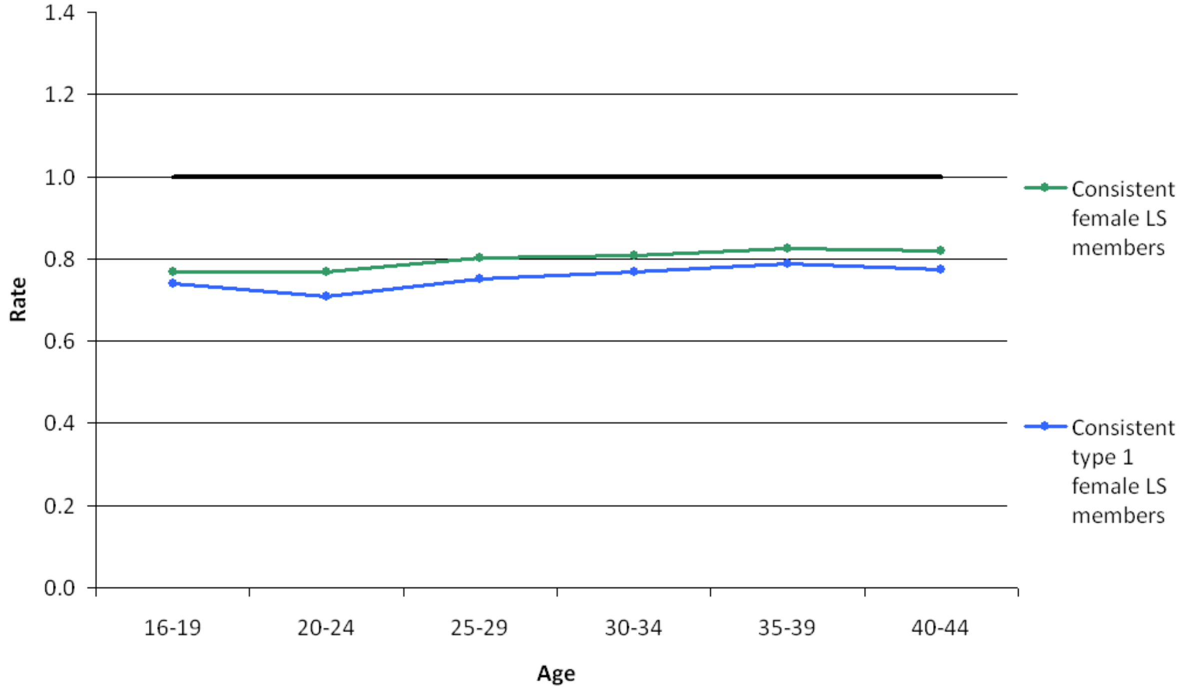


Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, Accessed June 2010.

Representation of LS births based on official statistics

As in the case of the sampling fraction of LS women based on official statistics, in the case of the representation of births based on official statistics there is a slightly lower, but less erratic, picture for all age groups. Figure 5.5 shows the 1991-2001 average. Generally, the trends in the representation fractions are similar to the sampling rates. However, there is less of a gap between the rates for the two categories used in the analysis. For the 1991-1996 period the two types are very closely matched. After 1996 there is an increasingly erratic trend at the end of the decade. In 1997-1999 the sampling fraction for the 25-29 year age group for the consistent Type 1 cases rises above the all consistent cases. At the end of the age range (40-44) in these years there is also a sizeable gap between the two types which is not observed earlier in the decade.

Figure 5.5: Consistent cases - representation of LS births based on official statistics, Average 1991-2001



Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, Accessed June 2010.

5.4.4 Conclusions

This section has worked only with the consistent cases in the LS. In much research using the LS it has been common to select a population matching that which is identified as consistent cases Type 1 – persons who are continuously resident for the entire decade between two census dates. However, given the work in Chapter 4 to establish the different forms of residence which LS members have, and taking into account the periods when they have embarked, it has been possible to run finer-grained analysis of what have been collectively identified as consistent cases in the LS. Interesting is the limited degree to which the two differ in their sampling and representation fractions. However, throughout the decade there has been an increasing difference between the two types. In the latter part of the decade the all consistent cases type is substantially closer to the 1.09% target.

In the comparison of the ASFRs there is a small difference between both samples and the official rate. However, the all consistent cases category is better at representing the fertility rates in younger age groups. This is important: the all consistent cases category includes entries to the LS in all the age groups, and these are highest among the younger ages as this is where there is more migration. For this reason the selection of a sample based on residence at both census dates at either end of a decade is unsatisfactory.

Comparisons with mid-year estimates for women in the LS show that there is a lower than anticipated sampling of women in the LS. Generally, the rates vary between 0.8 and 1 with the older ages closer to and, in some cases, above 1. In the latter part of the decade (i.e. years 1998-2001) there is a widening gap between the consistent Type 1 cases and the all consistent cases in the data. This again highlights that the inclusion of women through the more open all consistent cases type is desirable. Importantly, the difference between the two types is in the key childbearing age groups (20-24 years through to 30-34 years).

For births in the LS there is a more dramatic picture. Sampling fractions show a more pronounced difference between the age groups, with the 16-19 year age group having a rate around 0.7 and the 40+ age group a rate around 1% (for the decade on average). As with the rates for women in the LS, there is a widening gap between the Type 1 cases and the all consistent cases later in the decade. In the representation fractions there is less of a gap between the two categories, but the all consistent cases is still a better sample for analysis based on these results.

Overall, this section has identified that, although the Type 1 consistent cases category (those persons resident for the 1991-2001 period without any recorded embarkation or re-entry) gives rates and sampling fractions which are close to the FM1 ASFR and TFR figures, the all consistent cases type is a better sample, particularly in the later part of the decade. This is likely to be a result of the ability of the all consistent cases category to be able to include new entries to the LS and re-entries to the LS. In terms of the age effect of using this category, it seems that it is in the key childbearing ages 25-29 and 30-34 years where there is the biggest difference between the two categories and where the all consistent cases category is giving a better representation.

The next section works with the other broad category of inconsistent cases, which are more problematic. Consistent cases, as identified above, are added to the inconsistent cases to provide an adjusted sample which allocates some exposure to risk to the members of the LS who fall into an inconsistent type. As already hypothesised after the analysis of the sampling of women in the LS above, it is possible that the denominator might become inflated and this could reduce the ASFRs.

5.5 Using inconsistent cases in the LS - fertility rates, births and women

A large amount of work has focused on identifying the cases in the LS which are somehow incomplete, or contain some form of missing information. Chapter 4 fully outlined the different forms of incompleteness which can be identified and then accounted for all LS members in the 1991-2001 and 2001-2007 periods. For each decade LS members are fully coded for each of the different types of trajectory that they had in the relevant time period. For some of these cases where the incompleteness is relatively straightforward (i.e. one missing event), it is possible to allocate exposure to risk for these LS members. In the earlier part of this chapter (methodology), information was provided on the cases which have had some form of exposure to risk allocated to them, and those which have been excluded from analysis altogether. In this part of the work the fertility rates calculated using the inconsistent cases which have been able to allocate exposure in addition to the consistent cases identified, are reported. As in the last section, comparisons are made to the official statistics from the ONS.

5.5.1 Fertility rates

The ASFRs arrived at through using the all consistent and inconsistent cases are only slightly different from the consistent cases rates. However, there is an impact on the number of women included in the analysis – the denominator is increased substantially in the younger age groups, with over 2,669 extra women in 1991 for the 15-19 year age group and 2,690 extra women in the 20-24 year age group. These numbers are fairly typical for other years and the same age groups in the 1991-2001 period. Among the older age groups there is less of an impact on the denominator.

Tables 5.16-5.27 show the number of women for the all consistent and all consistent and inconsistent cases. They also show the official figures for each year as a comparison. Despite the increases in the number of women in some of the age groups, there are relatively minor changes to the fertility rates because the number of births also increases.

Table 5.16: Consistent and inconsistent cases - women, births and ASFRs for 1991

AGE	ALL CONSISTENT CASES			CONSISTENT AND INCONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	13,803	428	31.0	16,472	507	30.8	33.0
20-24	17,055	1,448	84.9	19,745	1,609	81.5	89.3
25-29	18,794	2,205	117.3	21,277	2,428	114.1	119.4
30-34	17,960	1,501	83.6	19,924	1,623	81.5	86.7
35-39	16,528	473	28.6	18,139	515	28.4	32.1
40-44	17,968	71	4.0	19,490	78	4.0	5.1
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.75			1.70	1.82

*Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.17: Consistent and inconsistent cases - women, births and ASFRs for 1992

AGE	ALL CONSISTENT CASES			CONSISTENT AND INCONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	13,101	390	29.8	15,939	470	29.5	31.7
20-24	16,504	1,430	86.6	19,532	1,630	83.5	86.2
25-29	18,770	2,115	112.7	21,479	2,342	109.0	117.5
30-34	18,385	1,566	85.2	20,480	1,699	83.0	87.3
35-39	16,803	519	30.9	18,491	567	30.7	33.4
40-44	17,200	96	5.6	18,755	111	5.9	5.5
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.75			1.71	1.80

*Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.18: Consistent and inconsistent cases - women, births and ASFRs for 1993

AGE	ALL CONSISTENT CASES			CONSISTENT AND INCONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	12,857	344	26.8	15,928	423	26.6	30.9
20-24	16,114	1,194	74.1	19,453	1,364	70.1	82.6
25-29	18,529	2,122	114.5	21,435	2,306	107.6	114.4
30-34	18,662	1,494	80.1	20,847	1,606	77.0	87.3
35-39	17,199	578	33.6	18,955	619	32.7	34.1
40-44	16,794	106	6.3	18,378	111	6.0	5.9
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.68			1.60	1.77

*Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.19: Consistent and inconsistent cases - women, births and ASFRs for 1994

AGE	ALL CONSISTENT CASES			CONSISTENT AND INCONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	12,924	357	27.6	16,257	432	26.6	28.9
20-24	15,644	1,131	72.3	19,293	1,321	68.5	79.1
25-29	18,426	1,996	108.3	21,487	2,192	102.0	112.4
30-34	18,973	1,528	80.5	21,266	1,677	78.9	89.3
35-39	17,513	569	32.5	19,329	627	32.4	35.8
40-44	16,501	118	7.2	18,118	127	7.0	6.1
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.64			1.58	1.75

*Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.20: Consistent and inconsistent cases - women, births and ASFRs for 1995

AGE	ALL CONSISTENT CASES			CONSISTENT AND INCONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	13,308	336	25.2	16,992	408	24.0	28.5
20-24	15,337	1,087	70.9	19,308	1,281	66.3	76.4
25-29	17,861	1,865	104.4	21,087	2,059	97.6	108.7
30-34	19,221	1,631	84.9	21,597	1,777	82.3	88.2
35-39	17,829	596	33.4	19,685	648	32.9	36.4
40-44	16,401	113	6.9	18,042	121	6.7	6.5
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.63			1.55	1.72

*Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.21: Consistent and inconsistent cases - women, births and ASFRs for 1996

AGE	ALL CONSISTENT CASES			CONSISTENT AND INCONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	13,714	374	27.3	19,126	490	25.6	29.7
20-24	14,384	1,153	80.2	20,858	1,420	68.1	77.0
25-29	17,629	1,830	103.8	23,057	2,226	96.5	106.9
30-34	19,102	1,699	88.9	23,043	1,996	86.6	89.7
35-39	18,085	650	35.9	21,219	759	35.8	37.5
40-44	16,555	117	7.1	19,428	141	7.3	6.9
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.72			1.60	1.74

*Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.22: Consistent and inconsistent cases - women, births and ASFRs for 1997

AGE	ALL CONSISTENT CASES			CONSISTENT AND INCONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	13,968	381	27.3	15,298	398	26.0	30.3
20-24	13,701	998	72.8	15,945	1,072	67.2	75.9
25-29	17,130	1,798	105.0	19,219	1,954	101.7	104.5
30-34	19,074	1,596	83.7	20,548	1,745	84.9	89.8
35-39	18,508	698	37.7	19,752	748	37.9	39.3
40-44	16,833	121	7.2	18,037	129	7.2	7.3
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.67			1.63	1.73

*Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.23: Consistent and inconsistent cases - women, births and ASFRs for 1998

AGE	ALL CONSISTENT CASES			CONSISTENT AND INCONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	13,913	437	31.4	15,246	466	30.6	31.2
20-24	13,503	995	73.7	15,746	1,057	67.1	74.9
25-29	16,804	1,621	96.5	18,894	1,765	93.4	101.6
30-34	18,897	1,614	85.4	20,372	1,779	87.3	90.7
35-39	18,808	695	37.0	20,052	747	37.3	40.4
40-44	17,230	124	7.2	18,433	132	7.2	7.5
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.66			1.62	1.73

*Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.24: Consistent and inconsistent cases - women, births and ASFRs for 1999

AGE	ALL CONSISTENT CASES			CONSISTENT AND INCONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	14,158	465	32.8	15,494	491	31.7	31.1
20-24	13,606	900	66.1	15,850	955	60.3	73.1
25-29	16,429	1,643	100.0	18,518	1,778	96.0	98.4
30-34	18,841	1,608	85.3	20,316	1,757	86.5	89.7
35-39	19,166	727	37.9	20,410	779	38.2	40.6
40-44	17,556	104	5.9	18,759	114	6.1	7.7
45-49	*	*	<0.5	*	*	<0.5	0.4
TFR			1.64			1.59	1.70

*Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.25: Consistent and inconsistent cases - women, births and ASFRs for 2000

AGE	ALL CONSISTENT CASES			CONSISTENT AND INCONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	14,490	347	23.9	15,829	364	23.0	29.5
20-24	14,042	887	63.2	16,287	939	57.7	70.2
25-29	16,206	1,466	90.5	18,294	1,581	86.4	94.5
30-34	18,298	1,577	86.2	19,772	1,749	88.5	88.1
35-39	19,469	734	37.7	20,712	800	38.6	41.4
40-44	17,893	114	6.4	19,097	123	6.4	8.0
45-49	*	*	<0.5	*	*	<0.5	0.4
TFR			1.54			1.50	1.66

*Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.26: Consistent and inconsistent cases - women, births and ASFRs for 2001

AGE	ALL CONSISTENT CASES			CONSISTENT AND INCONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	15,002	375	25.0	16,379	400	24.4	28.0
20-24	14,593	916	62.8	16,856	978	58.0	69.3
25-29	15,360	1,492	97.1	17,454	1,598	91.6	91.9
30-34	18,145	1,593	87.8	19,622	1,771	90.3	88.2
35-39	19,396	746	38.5	20,640	843	40.8	41.6
40-44	18,173	132	7.3	19,379	141	7.3	8.4
45-49	*	*	<0.5	*	*	<0.5	0.5
TFR			1.59			1.56	1.64

*Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011. Note - * denotes values deleted to meet ONS disclosure controls.*

Table 5.27: Consistent and inconsistent cases - women, births and ASFRs for 1991-2001

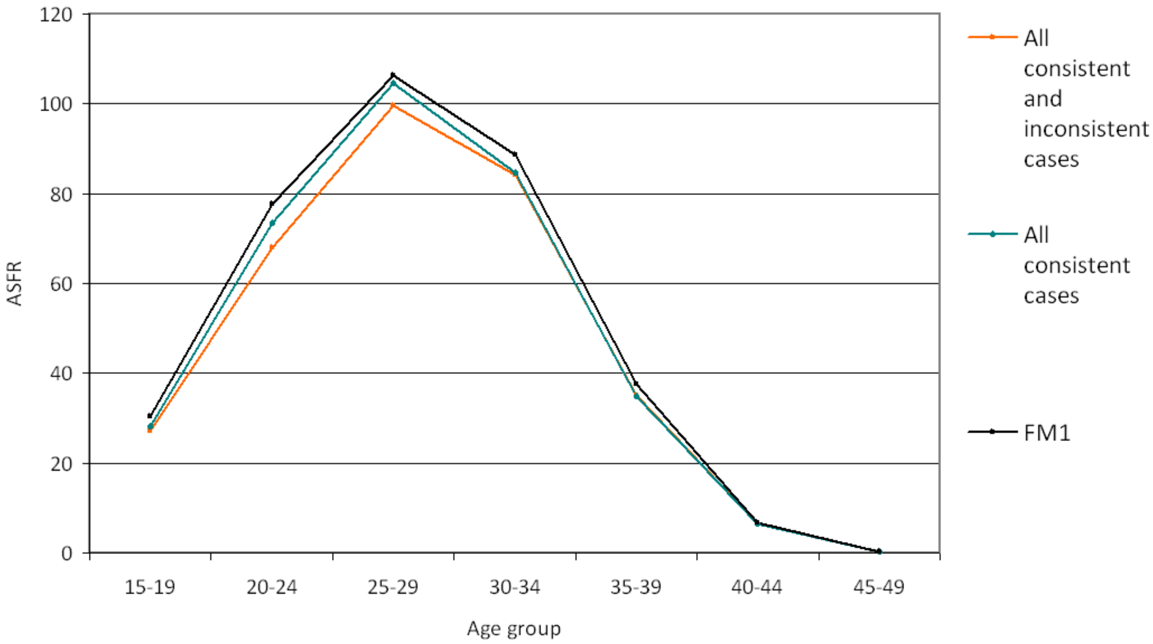
AGE	ALL CONSISTENT CASES			CONSISTENT AND INCONSISTENT CASES			FM 1 - OFFICIAL RATES
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Rate per 1,000
15-19	13,749	385	28.0	16,269	441	27.2	30.3
20-24	14,953	1,104	73.4	18,079	1,239	68.0	77.6
25-29	17,449	1,832	104.6	20,200	2,021	99.6	106.4
30-34	18,687	1,582	84.7	20,708	1,744	84.2	88.6
35-39	18,119	635	34.9	19,762	696	35.1	37.5
40-44	17,191	111	6.4	18,720	121	6.5	6.8
45-49	*	*	<0.5	*	*	<0.5	0.3
TFR			1.66			1.60	1.74

*Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011. Note - * denotes values deleted to meet ONS disclosure controls.*

Comparing fertility rates with Official Statistics rates

Figure 5.6 shows the ASFRs for the 1991 – 2001 period from the table above, with the inclusion of the FM1 – official statistics figure for comparison. This shows that there is a lower comparability with the official rate for the all consistent and inconsistent cases. In the 1992-2001 period the rates for the combined consistent and inconsistent cases are slightly lower than the official statistics and the consistent cases on their own. Although throughout the years the combined consistent and inconsistent cases show a lower rate of comparability with the official figures than just the consistent cases alone, the profile of the rates generally follows that of the official rates and the consistent cases ASFRs.

Figure 5.6: Consistent and inconsistent cases - age-specific fertility rates, average / overall 1991-2001



Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011.

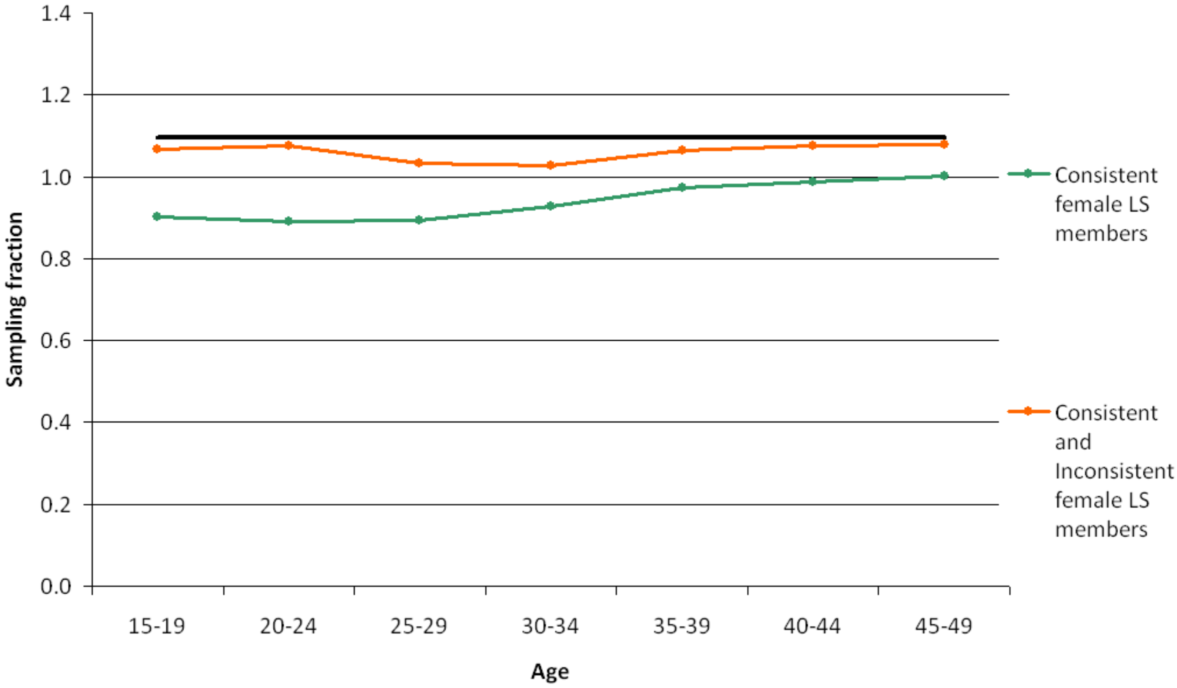
5.5.2 Comparing the denominator (women) with ONS mid-year estimates

Given the trends identified in Chapter 4, with certain age groups showing a lower representation in the LS compared with others – particularly where there is attrition in the study, the comparison with ONS mid-year estimates is especially important for inconsistent cases.

Sampling fraction of official statistics women by the LS

Figure 5.7 shows the average sampling fraction of official statistics women by the LS for the 1991-2001 period. The profile of the two categories is similar, although the combined consistent and inconsistent cases show a higher overall sampling fraction above 1% for 1991-1996 (see appendix) and a good match with the 1.09% target which is denoted by the thick black line. For the 1997-2001 period there is a lower rate of sampling in the 15-19 and 20-24 age groups than the 1.09% target, but the rates are close. At the older ages there is a higher rate of sampling than 1% in these years. This shows that the combined consistent and inconsistent cases are close to the 1.09% target and that the drop in sampling around the 25-34 age groups is consistent throughout the decade. Overall, the inclusion of the inconsistent cases in the sample for analysis has increased the sampling of women in the LS and led to sampling fractions closer to the 1.09% target.

Figure 5.7: Consistent and inconsistent cases - sampling fraction of official statistics women by the LS, Average 1991-2001

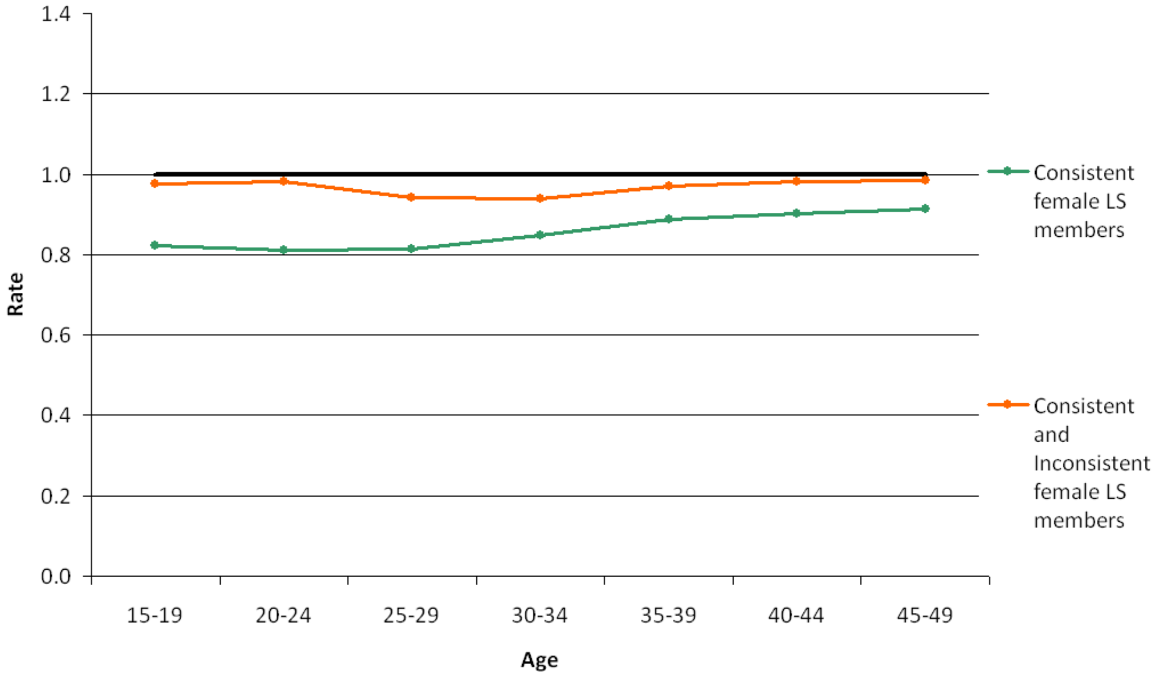


Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 Local Authority Population Studies: 07/10/04, Accessed June 2010.

Representation of LS women based on official statistics

As with the sampling fractions above, the rates of representation for LS births based on the 'expected' number show similar profiles for the two categories, but with higher rates for the combined consistent and inconsistent cases category than just the consistent cases category. For the years 1991-1995 the combined consistent and inconsistent categories are close to the target rate of 1%. As with the sampling fractions, in 1996 there are rates above 1% for all age groups. This must be related to the specific criteria applied to this year and the re-entry date used for the imputed re-entry for inconsistent cases (which was 1996). For the rest of the decade (1997-2001) the rates are below the target of 1% but at the older age groups show a higher degree of representation. Figure 5.8 shows the average figures for the decade and that the rates for all consistent cases and the consistent and inconsistent cases are comparable with the 1% target, but there is a lower rate for the 25-34 year old age group. It is clear there is a high degree of representation for women using the inconsistent cases in addition to the consistent cases. The fit of the rates with target as a result of the inclusion of the inconsistent women in the sample is stronger.

Figure 5.8: Consistent and inconsistent cases - representation of LS women based on official statistics, 1991-2001



Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 Local Authority Population Studies: 07/10/04, Accessed June 2010.

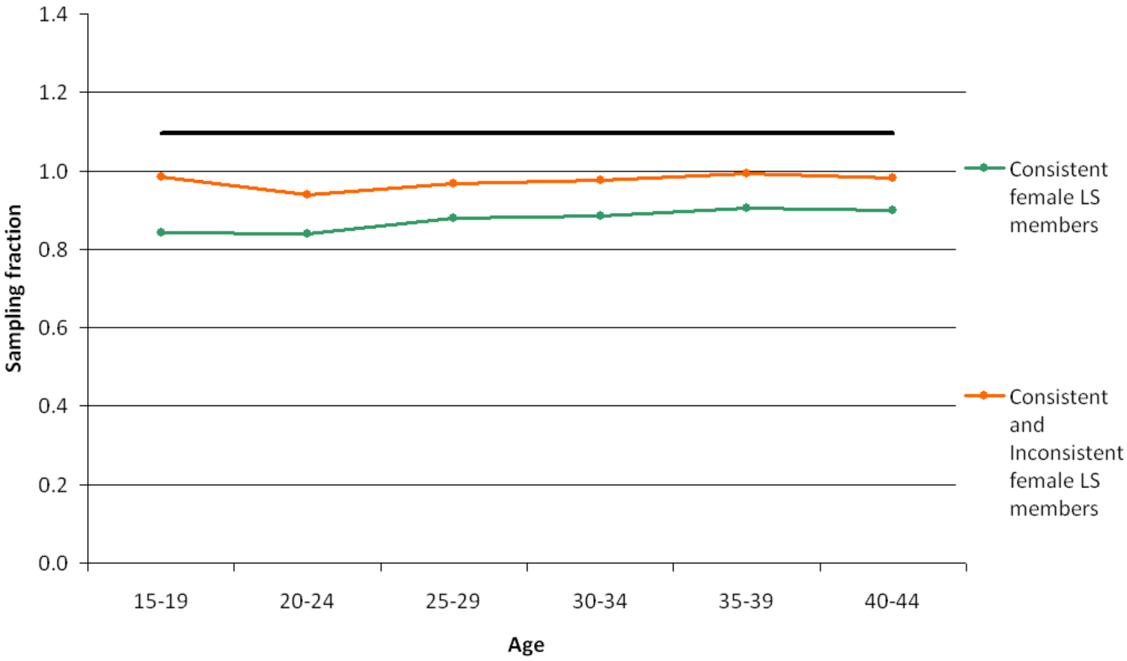
5.5.3 Comparing the numerator (births) with ONS vital statistics

This section uses both the consistent and inconsistent births from the LS to compare with the ONS official births statistics. Sampling fractions and the representation of LS births based on official statistics are calculated.

Sampling fraction of official statistics women by the LS

The sampling fractions for the all consistent cases births have already been detailed in this chapter. However, here they are compared to the combined consistent and inconsistent births. Generally, the combined consistent and inconsistent cases show a higher degree of comparability to the 1.09% target. Figure 5.9 shows the sampling fractions for the 1991-2001 period. In 1996 (see appendix) there is a large difference between the consistent and combined categories. As explained previously, this difference must be the result of the criteria used for the selection of the sample from the LS in this year. At the younger ages in the graphs (see appendix), the sampling fractions at the youngest ages are much higher than for just the consistent cases. This highlights the difference in the sample selected at the youngest age group when the consistent cases alone are used. On average there is a sampling fraction consistently around 1%. This is higher than for the consistent cases alone but still not matching the 1.09% target which is denoted by the thick black line.

Figure 5.9: Consistent and inconsistent cases - sampling fraction of official statistics women by the LS, Average 1991-2001

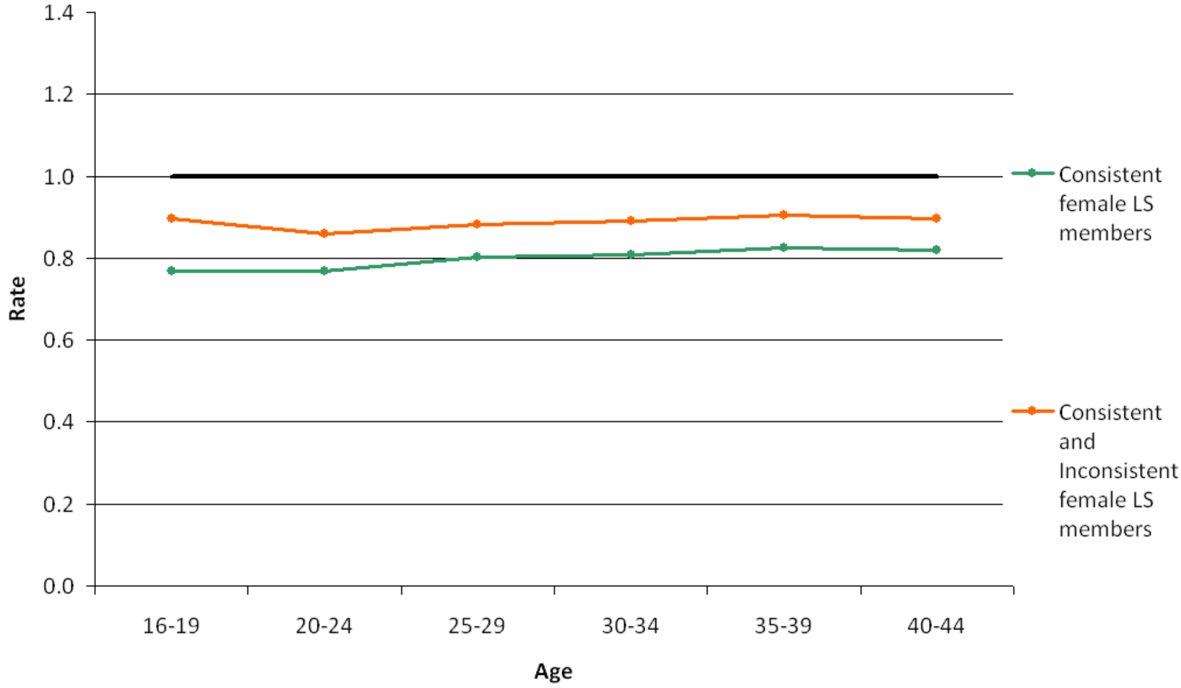


Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, Accessed June 2010.

Representation of LS births based on official statistics

Figures 5.10 shows the representation of births in the LS based on the corresponding official statistics figures. As with the sampling fractions for the 1991-2001 period, there is a higher rate for the 16-19 year age group than all the others. However, the most stable representation fraction across the whole time-scale is for the all consistent cases. This shows that there is the greatest consistency in the representation of births among the consistent cases even if the level is somewhat lower than the combined consistent and inconsistent cases.

Figure 5.10: Consistent and inconsistent cases - representation of LS births based on official statistics, Average 1991-2001



Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, Accessed June 2010.

5.5.4 Conclusions

This analysis has used both the consistent and inconsistent cases as identified in the last chapter. Some LS members have been dropped from this analysis – these are cases with too little complete information for them to be of use, or where the degree of 'missingness' or inconsistency is too high to be compensated for.

Overall, the ASFRs arrived at using the sample selected are comparable with the official statistics. Generally, the rates are lower than for the all consistent cases category. In many cases the profile of the ASFRs is also not as good as for the consistent cases alone. At older ages there is no identifiable difference between the all consistent cases and the consistent and inconsistent cases. It is among the 20-24, 25-29 and 30-34 age groups where there is the greatest difference between the consistent cases and the combined consistent and inconsistent cases.

For the sampling fractions of women using the combined consistent and inconsistent cases there is a higher rate than using the consistent cases alone. Figure 5.9 showed that the consistent and inconsistent cases are better at representing the population of women for the younger age groups compared to the consistent cases alone. In the key childbearing ages (25-34) there is a slightly lower level of sampling than for the other age groups.

It seems likely that the inclusion of the selected inconsistent cases has decreased the fertility rates because too much extra time exposed to risk of birth has been granted to persons who may not have been resident in England and Wales for all of the time. The corresponding births to this group do not make up for the increased exposure time granted to them. This may be a result of the imputation, but the use of these cases would not be possible without some form of imputation solution to make up for the missing information.

5.6 Foreign-born women in the LS - fertility rates, births and women

A key part of this research is concerned with the fertility of recent immigrants to England and Wales. Therefore, the fertility of foreign born women in the LS should be considered here. This section considers the representation of births to foreign born women in the LS. Migration to the UK increased from the late 1990s. Country of birth data for LS members is collected at the decennial census, and also for LS members who give birth, through vital registration data. It is possible to group the countries of birth so they correspond with those in the ONS FM1 fertility series and create one group of 'foreign born' women to give an overall comparison. Note that for the Other European Union category each year in the series matches the membership of the EU in that year (excluding UK and Ireland).

5.6.1 Percentage of births to women from selected countries in FM1 volume

No official age-specific or total fertility rates for the mother's country of birth are published by the ONS. This seems to be because it is difficult to identify the appropriate denominator. In the next section these statistics are calculated using those persons resident at a census. However, the best benchmark is the percentage of births to foreign born women compared to the total population. At birth the mother's country of birth is recorded and only the numerator and denominator from the LS is needed to make the comparison with the official statistics percentages. Tables 5.28 to 5.31 show the percentages of births to women from selected countries which are included in the annual FM1 volume. For using the LS to look at the fertility of recent migrants to England and Wales and the representation of the LS for foreign born mothers this is a vital exercise. In the analysis here both the 1991 and 2001 variables on country of birth of the mother (asked at the census) have been used. Analysis has been completed for all consistent cases and the consistent Type 1 cases.

Table 5.28 shows the percentage of births by country of birth of mother in the LS all consistent cases for the 1991-2001 period. Residence at the 1991 census is needed for inclusion in this sample, as otherwise there is no country of birth information for the LS member (this is collected at the census). For the UK-born the percentages from the LS are fairly comparable. In the case of India there is a higher percentage from the LS in the early part of the decade and then a decline after this to percentages which are lower and less comparable. This could be related to the small numbers of migrants at young ages – it is unlikely that there are large flows of women who arrived in the UK as children, were resident at the 1991 census and then continuously resident in the 1990s. Pakistan also has a widening divergence in the latter part of the time period. Using the 2001 country of birth variable, the LS member had to be at the 2001 census, the percentage (Table 5.29) is much higher for the latter years of the 1990s and indeed highly comparable to the FM1 values. This trend is also true for Bangladesh – the 1991 variable produces percentages which are comparable for the early part of the decade but after 1996 the 2001 census variable provides a closer comparison. In the case of East Africa and Other European Union there is a higher level of comparability through time but there is still a better representation using the 2001 census variable than the 1991 for births later in the decade.

Table 5.28: Percentage of births to women from selected countries of birth – ONS LS, All consistent cases, 1991 census variable and ONS FM1 volume, 1991-2001 (All rates include only those members at the 1991 census)

YEAR	United Kingdom		Irish Republic		Australia, Canada and New Zealand		India		Pakistan		Bangladesh		East Africa		Other European Union	
	FM1 volume	LS perfect cases	FM1 volume	LS perfect cases	FM1 volume	LS perfect cases	FM1 volume	LS perfect cases	FM1 volume	LS perfect cases	FM1 volume	LS perfect cases	FM1 volume	LS perfect cases	FM1 volume	LS perfect cases
1991	88.5	90.7	0.9	0.5	0.4	0.3	1.2	1.8	1.8	1.5	0.8	0.9	0.9	0.7	1.1	0.7
1992	88.1	89.7	0.8	0.5	0.4	0.3	1.1	1.3	1.9	1.5	0.8	0.9	0.9	0.7	1.1	0.6
1993	87.9	85.1	0.8	0.5	0.5	0.3	1.1	0.8	1.9	1.6	0.9	0.8	0.8	0.6	1.2	0.7
1994	87.6	88.9	0.8	0.5	0.5	0.3	1.1	1.2	1.9	1.5	0.9	1.2	0.8	0.7	1.2	0.9
1995	87.4	88.4	0.8	0.3	0.5	0.2	1.0	0.9	1.9	1.1	1.0	1.0	0.8	0.4	1.4	0.9
1996	87.2	88.4	0.8	0.6	0.5	0.4	1.0	0.8	1.9	1.2	1.1	0.9	0.8	0.6	1.4	0.7
1997	86.9	87.8	0.8	0.4	0.5	0.3	1.0	0.5	2.0	1.1	1.1	1.0	0.7	0.4	1.5	0.6
1998	86.4	87.9	0.7	0.4	0.5	0.2	1.0	0.4	2.1	1.0	1.2	0.9	0.7	0.5	1.6	0.6
1999	85.7	87.0	0.7	0.2	0.6	0.2	1.0	0.3	2.2	1.0	1.2	0.7	0.7	0.4	1.7	0.9
2000	84.5	84.5	0.7	0.3	0.6	0.1	1.1	0.4	2.2	0.9	1.2	1.0	0.7	0.3	1.8	0.7
2001	83.5	84.5	0.6	0.3	0.6	0.3	1.1	0.4	2.5	0.9	1.4	0.7	0.6	0.4	1.8	0.6

Own elaboration based on ONS LS, September 2010, FM1 no.23, 1994, FM1 no.27, 1998 FM1 no.32, 2003.

Table 5.29: Percentage of births to women from selected countries of birth – ONS LS, All consistent cases, 2001 census variable and ONS FM1 volume, 1991-2001 (All rates include only those members at the 2001 census)

YEAR	United Kingdom		Irish Republic		Australia, Canada and New Zealand		India		Pakistan		Bangladesh		East Africa		Other European Union	
	FM1 volume	LS perfect cases	FM1 volume	LS perfect cases	FM1 volume	LS perfect cases	FM1 volume	LS perfect cases	FM1 volume	LS perfect cases	FM1 volume	LS perfect cases	FM1 volume	LS perfect cases	FM1 volume	LS perfect cases
1991	88.5	89.5	0.9	0.5	0.4	0.3	1.2	1.7	1.8	1.5	0.8	0.9	0.9	0.6	1.1	0.8
1992	88.1	89.0	0.8	0.5	0.4	0.3	1.1	1.2	1.9	1.6	0.8	1.2	0.9	0.7	1.1	0.6
1993	87.9	84.6	0.8	0.4	0.5	0.3	1.1	1.1	1.9	1.6	0.9	0.9	0.8	0.6	1.2	0.8
1994	87.6	88.3	0.8	0.5	0.5	0.3	1.1	1.5	1.9	1.8	0.9	1.6	0.8	0.7	1.2	1.1
1995	87.4	88.8	0.8	0.2	0.5	0.2	1.0	1.4	1.9	1.7	1.0	1.3	0.8	0.6	1.4	1.0
1996	87.2	88.5	0.8	0.6	0.5	0.5	1.0	1.3	1.9	1.6	1.1	1.3	0.8	0.7	1.4	0.9
1997	86.9	87.7	0.8	0.5	0.5	0.4	1.0	1.2	2.0	1.9	1.1	1.5	0.7	0.5	1.5	1.0
1998	86.4	88.3	0.7	0.5	0.5	0.3	1.0	1.1	2.1	1.9	1.2	1.4	0.7	0.6	1.6	1.0
1999	85.7	87.3	0.7	0.3	0.6	0.4	1.0	1.3	2.2	2.1	1.2	1.3	0.7	0.6	1.7	1.5
2000	84.5	85.1	0.7	0.5	0.6	0.4	1.1	1.7	2.2	2.1	1.2	1.8	0.7	0.5	1.8	1.5
2001	83.5	85.3	0.6	0.5	0.6	0.7	1.1	1.5	2.5	2.3	1.4	1.5	0.6	0.8	1.8	1.5

Own elaboration based on ONS LS, September 2010, FM1 no.23, 1994, FM1 no.27, 1998 FM1 no.32, 2003.

Table 5.30 and 5.31 show the same statistics but use the consistent Type 1 cases only. Again, the 1991 and 2001 census variables on country of birth have been used in the selection of the births. This means that for all the years in the table the LS member must have been resident at either the 1991 census where the country of birth variable has been used or the 2001 census where the country of birth variable has been used. In Table 5.30 the percentages of births to UK-born women are much higher than in the last two analyses – in all the years they are over 91%. For the other countries in this table the percentages are lower than the corresponding FM1 figures – this is especially the case for India, Pakistan, East Africa and the Other European Union category. In the case of Table 5.31 the UK figures show a higher percentage than the corresponding FM1 figures. This must be a reflection of the characteristics of the sample to an extent – the percentage of LS members who are a Type 1 case is highest among the UK born group. Interestingly, there is not the same trend with regard to the changes in percentages at the years later in the series (explained above). This is because these LS members would have been at the 1991 and 2001 censuses.

Table 5.30: Percentage of births to women from selected countries of birth – ONS LS, All Type 1 consistent cases, 1991 census variable and ONS FM1, 1991-2001

YEAR	United Kingdom		Irish Republic		Australia, Canada and New Zealand		India		Pakistan		Bangladesh		East Africa		Other European Union	
	FM1 volume	LS type 1 perfect cases	FM1 volume	LS type 1 perfect cases	FM1 volume	LS type 1 perfect cases	FM1 volume	LS type 1 perfect cases	FM1 volume	LS type 1 perfect cases	FM1 volume	LS type 1 perfect cases	FM1 volume	LS type 1 perfect cases	FM1 volume	LS type 1 perfect cases
1991	88.5	91.1	0.9	0.4	0.4	0.3	1.2	1.8	1.8	1.5	0.8	0.9	0.9	0.6	1.1	0.7
1992	88.1	91.4	0.8	0.5	0.4	0.3	1.1	1.2	1.9	1.5	0.8	1.0	0.9	0.7	1.1	0.6
1993	87.9	89.2	0.8	0.4	0.5	0.3	1.1	0.8	1.9	1.6	0.9	0.8	0.8	0.6	1.1	0.7
1994	87.6	91.2	0.8	0.5	0.5	0.3	1.1	1.2	1.9	1.5	0.9	1.2	0.8	0.7	1.2	1.0
1995	87.4	92.7	0.8	0.2	0.5	0.2	1.0	0.9	1.9	1.2	1.0	1.0	0.8	0.5	1.4	0.9
1996	87.2	92.7	0.8	0.5	0.5	0.4	1.0	0.8	1.9	1.2	1.1	1.0	0.8	0.6	1.4	0.7
1997	86.9	93.2	0.8	0.5	0.5	0.3	1.0	0.6	2.0	1.1	1.1	1.0	0.7	0.4	1.5	0.7
1998	86.4	94.3	0.7	0.3	0.5	0.2	1.0	0.5	2.1	1.1	1.2	0.9	0.7	0.5	1.6	0.6
1999	85.7	94.5	0.7	0.3	0.6	0.2	1.0	0.3	2.2	1.0	1.2	0.8	0.7	0.4	1.7	0.9
2000	84.5	94.3	0.7	0.3	0.6	0.1	1.1	0.5	2.2	1.0	1.2	1.0	0.7	0.3	1.8	0.7
2001	83.5	94.9	0.6	0.3	0.6	0.3	1.1	0.4	2.5	0.9	1.4	0.8	0.6	0.5	1.8	0.7

Own elaboration based on ONS LS, September 2010, FM1 no.23, 1994, FM1 no.27, 1998 FM1 no.32, 2003.

Table 5.31: Percentage of births to women from selected countries of birth – ONS LS, All Type 1 consistent cases, 2001 census variable and ONS FM1, 1991-2001

YEAR	United Kingdom		Irish Republic		Australia, Canada and New Zealand		India		Pakistan		Bangladesh		East Africa		Other European Union	
	FM1 volume	LS type 1 perfect cases	FM1 volume	LS type 1 perfect cases	FM1 volume	LS type 1 perfect cases	FM1 volume	LS type 1 perfect cases	FM1 volume	LS type 1 perfect cases	FM1 volume	LS type 1 perfect cases	FM1 volume	LS type 1 perfect cases	FM1 volume	LS type 1 perfect cases
1991	88.5	91.1	0.9	0.5	0.4	0.3	1.2	1.7	1.8	1.5	0.8	0.8	0.9	0.6	1.1	0.8
1992	88.1	91.4	0.8	0.5	0.4	0.3	1.1	1.1	1.9	1.4	0.8	1.0	0.9	0.7	1.1	0.6
1993	87.9	89.0	0.8	0.4	0.5	0.3	1.1	0.9	1.9	1.5	0.9	0.8	0.8	0.5	1.2	0.8
1994	87.6	90.8	0.8	0.5	0.5	0.3	1.1	1.3	1.9	1.4	0.9	1.2	0.8	0.6	1.2	1.0
1995	87.4	92.7	0.8	0.2	0.5	0.2	1.0	0.8	1.9	1.2	1.0	0.9	0.8	0.5	1.4	0.8
1996	87.2	92.4	0.8	0.5	0.5	0.4	1.0	0.7	1.9	1.2	1.1	1.0	0.8	0.6	1.4	0.7
1997	86.9	93.0	0.8	0.5	0.5	0.3	1.0	0.5	2.0	1.2	1.1	1.0	0.7	0.5	1.5	0.8
1998	86.4	94.1	0.7	0.4	0.5	0.2	1.0	0.5	2.1	1.0	1.2	0.9	0.7	0.5	1.6	0.7
1999	85.7	94.2	0.7	0.2	0.6	0.2	1.0	0.3	2.2	1.1	1.2	0.8	0.7	0.4	1.7	0.9
2000	84.5	93.8	0.7	0.4	0.6	0.2	1.1	0.5	2.2	0.9	1.2	1.0	0.7	0.4	1.8	0.8
2001	83.5	94.6	0.6	0.3	0.6	0.3	1.1	0.4	2.5	0.9	1.4	0.8	0.6	0.6	1.8	0.7

Own elaboration based on ONS LS, September 2010, FM1 no.23, 1994, FM1 no.27, 1998 FM1 no.32, 2003.

5.6.2 Fertility rates

Tables 5.32 – 5.42 show the age-specific fertility rates for the same countries of birth in the tables above and use the 1991 census variable for country of birth of the LS member. As already discussed, there is no official statistics comparator. Those born elsewhere in the European Union show the highest degree of comparability with the UK. India and Bangladesh both show early peaks in their fertility rates – the highest ASFR is in the 20-24 year age group for these countries. Throughout the time series India and East Africa tend not to have births in the youngest age groups – shown by there not being any rates.

Tables 5.43 – 5.53 have the same information, but using the 2001 census country of birth variable. As with the 1991 census variable, there are some dramatic age-specific trends. Notable are those for India, Pakistan and Bangladesh. For these countries the rates are very high in the key childbearing ages (20-39 year) and then high even in the 40-44 year age group in the case of Bangladesh. By including only those resident as of the 2001 census, these tables are capturing new entrants to the LS, and therefore migrants to England and Wales and their subsequent fertility. The Other European Union figures vary quite substantially – the teenage rates are high in some years (e.g. 1994, 1999) and low in most of the others. There is also a later peak age-specific rate in the Other European Union group than the UK born women. It is in the 30-34 year age group where the highest ASFR can be found, compared to the 25-29 year age group for UK born women.

Table 5.32: Age-specific and total fertility rates for 1991 by selected countries of birth – ONS LS, all consistent cases, 1991 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	25.2	0.0	65.2	0.0	33.3	15.2	0.0	24.1
20-24	76.9	67.1	60.6	146.8	74.1	171.9	28.6	58.3
25-29	111.8	66.3	44.2	184.4	171.9	150.0	36.1	105.8
30-34	87.6	95.2	72.9	98.1	173.9	200.0	84.0	56.3
35-39	31.7	43.2	55.6	18.6	41.8	41.2	55.8	35.0
40-44	4.5	0.0	0.0	7.0	11.7	43.5	18.6	0.0
45-49	0.2	3.0	0.0	0.0	0.0	37.7	0.0	0.0
TFR	1.69	1.37	1.49	2.27	2.53	3.30	1.12	1.40

Own elaboration based on ONS LS, September 2010.

Table 5.33: Age-specific and total fertility rates for 1992 by selected countries of birth – ONS LS, all consistent cases, 1991 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	24.5	57.1	0.0	0.0	20.8	0.0	0.0	0.0
20-24	76.7	67.8	53.3	131.3	161.8	178.3	22.0	76.9
25-29	107.7	52.9	44.1	116.5	172.2	196.1	135.3	75.5
30-34	87.0	122.4	115.8	73.3	139.5	141.0	70.8	62.5
35-39	33.2	44.4	13.2	37.3	49.4	100.0	40.2	22.1
40-44	5.8	4.0	0.0	2.4	25.4	61.2	5.8	13.2
45-49	0.4	0.0	0.0	0.0	0.0	19.2	0.0	0.0
TFR	1.68	1.74	1.13	1.80	2.85	3.48	1.37	1.25

Own elaboration based on ONS LS, September 2010.

Table 5.34: Age-specific and total fertility rates for 1993 by selected countries of birth – ONS LS, all consistent cases, 1991 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	21.3	0.0	0.0	0.0	11.6	25.0	0.0	11.6
20-24	67.0	54.9	86.2	63.3	139.3	192.6	12.0	93.8
25-29	105.8	78.5	56.7	100.0	187.2	114.3	80.5	97.6
30-34	84.2	79.5	61.2	37.2	147.1	37.5	105.5	94.0
35-39	35.5	53.6	35.7	38.9	80.3	56.1	40.7	15.6
40-44	6.0	12.8	14.7	5.0	4.7	108.7	5.6	0.0
45-49	0.2	0.0	0.0	0.0	7.2	0.0	0.0	0.0
TFR	1.60	1.40	1.27	1.22	2.89	2.67	1.22	1.56

Own elaboration based on ONS LS, September 2010.

Table 5.35: Age-specific and total fertility rates for 1994 by selected countries of birth – ONS LS, all consistent cases, 1991 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	21.3	0.0	0.0	0.0	23.8	27.0	0.0	33.7
20-24	63.0	45.5	22.7	111.1	213.0	192.6	57.1	65.9
25-29	101.4	58.5	67.2	133.0	179.5	211.0	71.4	119.4
30-34	80.5	105.3	29.4	67.8	120.0	206.9	93.8	95.2
35-39	35.6	32.1	79.5	33.1	54.7	68.6	35.9	43.1
40-44	7.9	8.6	0.0	9.3	13.5	20.4	5.0	14.2
45-49	0.5	0.0	0.0	0.0	6.5	0.0	0.0	0.0
TFR	1.55	1.25	0.99	1.77	3.06	3.63	1.32	1.86

Own elaboration based on ONS LS, September 2010.

Table 5.36: Age-specific and total fertility rates for 1995 by selected countries of birth – ONS LS, all consistent cases, 1991 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	19.5	0.0	0.0	0.0	13.5	61.9	0.0	10.1
20-24	63.5	19.2	20.4	60.0	107.5	187.5	34.5	57.5
25-29	94.6	74.7	25.9	99.3	147.5	118.2	39.7	105.5
30-34	84.0	39.0	66.0	88.8	115.7	123.7	72.2	118.1
35-39	35.1	13.0	22.2	28.8	47.6	88.9	20.2	17.5
40-44	7.3	9.7	0.0	4.3	4.2	21.3	9.4	13.5
45-49	0.1	3.5	0.0	0.0	6.3	0.0	0.0	4.6
TFR	1.52	0.80	0.67	1.41	2.21	3.01	0.88	1.63

Own elaboration based on ONS LS, September 2010.

Table 5.37: Age-specific and total fertility rates for 1996 by selected countries of birth – ONS LS, all consistent cases, 1991 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	24.6	33.3	0.0	0.0	69.0	34.5	0.0	21.1
20-24	65.9	153.8	66.7	48.8	87.9	210.5	134.6	82.4
25-29	95.5	67.6	80.8	123.9	135.8	165.4	75.5	97.1
30-34	87.8	93.6	85.5	87.0	145.5	140.0	90.9	58.1
35-39	37.3	55.2	65.9	25.9	47.6	87.5	42.4	26.3
40-44	7.9	11.2	14.3	4.2	15.2	20.8	4.3	14.3
45-49	0.2	0.0	0.0	0.0	5.9	10.9	0.0	0.0
TFR	1.60	2.07	1.57	1.45	2.53	3.35	1.74	1.50

Own elaboration based on ONS LS, September 2010.

Table 5.38: Age-specific and total fertility rates for 1997 by selected countries of birth – ONS LS, all consistent cases, 1991 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	24.1	0.0	33.3	0.0	24.4	25.6	0.0	21.1
20-24	63.2	88.2	23.8	34.5	94.7	183.2	47.6	81.4
25-29	93.0	52.6	42.9	60.6	132.4	224.8	44.0	71.0
30-34	82.2	62.1	30.3	71.4	115.4	100.0	71.9	75.9
35-39	40.3	77.5	98.9	26.2	66.1	64.1	23.8	8.5
40-44	8.4	17.3	0.0	0.0	15.3	20.2	4.0	14.9
45-49	0.2	4.0	0.0	0.0	0.0	0.0	0.0	5.2
TFR	1.56	1.51	1.15	0.96	2.24	3.09	0.96	1.39

Own elaboration based on ONS LS, September 2010.

Table 5.39: Age-specific and total fertility rates for 1998 by selected countries of birth – ONS LS, all consistent cases, 1991 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	26.7	0.0	62.5	0.0	58.8	0.0	0.0	23.0
20-24	64.2	62.5	0.0	0.0	58.8	148.8	90.9	89.9
25-29	87.4	92.0	55.6	50.0	163.9	169.1	61.0	58.8
30-34	81.1	67.8	60.6	82.2	133.7	114.3	88.4	81.6
35-39	39.5	34.5	21.1	12.5	54.6	88.6	27.8	8.1
40-44	7.8	12.1	12.2	4.1	11.0	28.0	4.1	16.4
45-49	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
TFR	1.53	1.34	1.06	0.74	2.40	2.74	1.36	1.39

Own elaboration based on ONS LS, September 2010.

Table 5.40: Age-specific and total fertility rates for 1999 by selected countries of birth – ONS LS, all consistent cases, 1991 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	27.3	20.0	0.0	0.0	0.0	0.0	0.0	37.0
20-24	60.7	35.7	55.6	0.0	132.5	125.0	0.0	65.2
25-29	87.1	31.7	23.3	31.7	148.1	160.6	57.1	108.7
30-34	82.9	46.0	55.1	47.9	139.2	110.1	92.9	106.0
35-39	39.7	33.8	51.5	21.8	53.3	34.9	26.5	50.8
40-44	7.4	0.0	11.6	2.1	7.3	29.4	4.0	9.0
45-49	0.2	0.0	0.0	2.4	4.5	0.0	0.0	0.0
TFR	1.53	0.84	0.99	0.53	2.43	2.30	0.90	1.88

Own elaboration based on ONS LS, September 2010.

Table 5.41: Age-specific and total fertility rates for 2000 by selected countries of birth – ONS LS, all consistent cases, 1991 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	21.0	19.2	24.4	0.0	38.5	19.6	0.0	23.8
20-24	54.0	38.5	0.0	0.0	68.5	173.5	0.0	70.7
25-29	78.3	60.0	81.6	120.0	150.5	178.1	87.7	64.5
30-34	80.3	106.3	27.0	64.3	87.9	127.3	55.6	65.4
35-39	39.1	20.7	19.4	28.7	46.3	72.9	11.3	55.0
40-44	7.1	6.7	0.0	2.2	22.1	0.0	20.4	17.7
45-49	0.4	5.0	0.0	0.0	4.2	0.0	0.0	0.0
TFR	1.40	1.28	0.76	1.08	2.09	2.86	0.87	1.49

Own elaboration based on ONS LS, September 2010.

Table 5.42: Age-specific and total fertility rates for 2001 by selected countries of birth – ONS LS, all consistent cases, 1991 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	20.8	19.6	0.0	0.0	37.0	22.2	0.0	0.0
20-24	61.3	33.3	0.0	0.0	155.2	69.0	0.0	63.2
25-29	82.4	51.3	44.4	97.6	202.2	155.6	78.4	104.8
30-34	82.3	72.5	61.9	90.1	104.9	93.8	122.6	69.2
35-39	39.0	30.5	55.6	28.0	27.3	40.4	43.5	41.3
40-44	8.3	14.3	21.7	4.7	23.9	51.3	4.3	8.6
45-49	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0
TFR	1.47	1.11	0.92	1.10	2.75	2.16	1.24	1.44

Own elaboration based on ONS LS, September 2010.

Table 5.43: Age-specific and total fertility rates for 1991 by selected countries of birth – ONS LS, all consistent cases, 2001 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	25.8	0.0	20.0	0.0	31.6	35.1	0.0	23.3
20-24	76.6	44.0	78.1	169.5	113.9	172.4	39.2	78.9
25-29	109.7	74.1	59.7	149.5	139.7	145.6	87.0	80.9
30-34	85.7	83.3	78.9	88.1	179.0	129.4	93.6	100.7
35-39	31.3	34.5	39.2	24.1	31.1	41.2	49.3	45.8
40-44	4.5	0.0	23.8	4.9	6.2	48.8	15.6	5.1
45-49	0.2	3.2	0.0	0.0	0.0	20.4	0.0	0.0
TFR	1.67	1.20	1.50	2.18	2.51	2.96	1.42	1.67

Own elaboration based on ONS LS, September 2010.

Table 5.44: Age-specific and total fertility rates for 1992 by selected countries of birth – ONS LS, all consistent cases, 2001 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	25.7	33.3	0.0	0.0	29.4	27.8	0.0	0.0
20-24	80.2	34.5	78.1	123.9	169.4	192.5	67.4	70.2
25-29	106.1	55.6	68.5	95.8	150.2	241.1	115.9	83.3
30-34	87.5	162.2	90.9	69.8	138.7	192.8	91.4	74.3
35-39	33.3	49.3	19.2	30.8	80.5	89.1	29.3	20.1
40-44	5.4	4.0	0.0	5.0	21.9	65.9	6.8	17.0
45-49	0.3	0.0	0.0	0.0	0.0	39.2	0.0	0.0
TFR	1.69	1.69	1.28	1.63	2.95	4.24	1.55	1.33

Own elaboration based on ONS LS, September 2010.

Table 5.45: Age-specific and total fertility rates for 1993 by selected countries of birth – ONS LS, all consistent cases, 2001 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	22.1	0.0	0.0	30.3	30.0	18.2	0.0	29.7
20-24	70.9	13.3	82.0	85.5	152.9	201.2	23.3	80.0
25-29	107.2	69.6	60.2	126.1	150.9	189.4	98.7	94.7
30-34	84.6	101.9	60.2	37.6	125.5	50.0	88.7	113.3
35-39	35.5	66.2	47.6	34.7	84.3	54.5	34.3	40.5
40-44	5.7	4.6	0.0	5.3	4.8	83.3	6.3	0.0
45-49	0.5	0.0	0.0	0.0	0.0	17.9	0.0	0.0
TFR	1.63	1.28	1.25	1.60	2.74	3.07	1.26	1.79

Own elaboration based on ONS LS, September 2010.

Table 5.46: Age-specific and total fertility rates for 1994 by selected countries of birth – ONS LS, all consistent cases, 2001 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	20.8	0.0	0.0	0.0	61.9	80.0	0.0	37.4
20-24	65.5	79.4	64.5	137.3	204.8	260.1	53.3	58.4
25-29	101.7	78.7	104.7	156.3	154.5	216.8	77.5	82.4
30-34	80.6	122.4	59.5	69.2	105.7	206.9	84.2	125.7
35-39	35.9	44.1	69.4	35.5	54.1	66.0	33.3	75.3
40-44	8.0	4.7	0.0	12.2	22.5	10.4	5.6	11.4
45-49	0.5	0.0	0.0	0.0	6.4	0.0	0.0	0.0
TFR	1.57	1.65	1.49	2.05	3.05	4.20	1.27	1.95

Own elaboration based on ONS LS, September 2010.

Table 5.47: Age-specific and total fertility rates for 1995 by selected countries of birth – ONS LS, all consistent cases, 2001 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	19.9	0.0	0.0	0.0	43.5	72.9	0.0	16.3
20-24	66.2	0.0	15.4	180.0	157.2	214.7	15.6	63.7
25-29	97.9	56.5	56.8	141.3	159.0	157.9	58.0	69.0
30-34	85.9	60.0	94.1	100.0	98.0	100.0	123.6	100.6
35-39	36.4	30.1	52.6	23.9	63.2	53.2	41.5	62.5
40-44	7.6	0.0	0.0	6.7	16.4	28.8	16.0	11.2
45-49	0.1	3.6	0.0	0.0	6.6	0.0	0.0	4.6
TFR	1.57	0.75	1.09	2.26	2.72	3.14	1.27	1.64

Own elaboration based on ONS LS, September 2010.

Table 5.48: Age-specific and total fertility rates for 1996 by selected countries of birth – ONS LS, all consistent cases, 2001 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	24.3	35.7	0.0	28.6	82.4	11.1	0.0	16.7
20-24	69.7	116.7	98.4	104.2	146.5	220.1	100.0	44.0
25-29	99.4	74.1	107.5	192.1	165.3	154.3	93.8	55.0
30-34	90.2	121.7	107.1	89.6	133.1	136.4	125.0	98.4
35-39	38.0	65.6	92.0	19.9	33.1	90.9	51.6	52.9
40-44	7.8	13.3	0.0	4.3	26.3	9.8	4.9	17.6
45-49	0.2	0.0	0.0	0.0	6.0	11.5	0.0	0.0
TFR	1.65	2.14	2.02	2.19	2.96	3.17	1.88	1.42

Own elaboration based on ONS LS, September 2010.

Table 5.49: Age-specific and total fertility rates for 1997 by selected countries of birth – ONS LS, all consistent cases, 2001 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	23.4	0.0	29.4	0.0	14.7	23.0	0.0	16.5
20-24	64.7	37.0	18.2	210.5	185.2	253.3	37.7	55.9
25-29	95.8	74.8	60.6	120.9	162.2	195.5	62.5	64.4
30-34	84.7	93.2	44.4	74.3	100.8	105.7	109.2	83.3
35-39	41.0	105.3	95.7	30.0	94.1	34.1	33.7	34.3
40-44	8.3	6.9	0.0	0.0	18.7	19.4	4.8	23.1
45-49	0.2	4.0	0.0	0.0	0.0	0.0	0.0	0.0
TFR	1.59	1.61	1.24	2.18	2.88	3.16	1.24	1.39

Own elaboration based on ONS LS, September 2010.

Table 5.50: Age-specific and total fertility rates for 1998 by selected countries of birth – ONS LS, all consistent cases, 2001 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	26.7	0.0	25.0	62.5	90.9	0.0	0.0	25.6
20-24	66.8	69.0	0.0	174.4	166.7	200.0	90.9	46.6
25-29	92.5	68.6	37.4	95.2	190.7	153.0	64.2	49.1
30-34	84.1	104.8	109.1	95.2	120.0	118.9	98.8	91.6
35-39	41.4	52.6	43.0	18.1	82.0	81.4	33.2	33.3
40-44	8.2	7.3	0.0	4.2	11.2	44.6	9.5	16.9
45-49	0.1	0.0	0.0	0.0	4.7	0.0	0.0	0.0
TFR	1.60	1.51	1.07	2.25	3.33	2.99	1.48	1.32

Own elaboration based on ONS LS, September 2010.

Table 5.51: Age-specific and total fertility rates for 1999 by selected countries of birth – ONS LS, all consistent cases, 2001 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	26.0	24.4	0.0	52.6	55.6	15.4	31.3	37.4
20-24	61.7	31.7	0.0	152.2	196.5	201.3	0.0	55.3
25-29	92.5	51.0	42.7	134.4	201.8	147.4	79.2	60.8
30-34	84.9	50.0	90.9	83.0	131.3	125.8	103.2	120.2
35-39	40.6	38.5	82.5	19.1	82.0	43.5	45.7	66.7
40-44	7.6	0.0	39.0	4.3	7.5	37.0	4.6	22.9
45-49	0.3	0.0	0.0	0.0	4.4	0.0	0.0	0.0
TFR	1.57	0.98	1.28	2.23	3.40	2.85	1.32	1.82

Own elaboration based on ONS LS, September 2010.

Table 5.52: Age-specific and total fertility rates for 2000 by selected countries of birth – ONS LS, all consistent cases, 2001 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	17.9	19.2	24.4	0.0	153.8	19.6	0.0	25.6
20-24	50.4	0.0	115.4	904.8	397.3	336.7	0.0	47.0
25-29	76.9	100.0	163.3	700.0	408.6	294.5	175.4	48.5
30-34	82.9	106.3	72.1	150.0	159.3	245.5	103.2	111.5
35-39	40.1	34.5	38.8	37.4	60.2	93.8	16.9	49.3
40-44	7.0	6.7	22.5	4.5	14.7	0.0	16.3	11.0
45-49	0.4	0.0	0.0	2.2	0.0	0.0	0.0	0.0
TFR	1.38	1.33	2.18	8.99	5.97	4.95	1.56	1.46

Own elaboration based on ONS LS, September 2010.

Table 5.53: Age-specific and total fertility rates for 2001 by selected countries of birth – ONS LS, all consistent cases, 2001 census variable

AGE	United Kingdom	Irish Republic	Australia, Canada and New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union
15-19	19.7	0.0	0.0	0.0	54.5	67.8	25.6	16.3
20-24	61.2	25.6	31.3	193.2	276.2	157.5	0.0	30.5
25-29	86.4	31.9	50.8	160.2	201.0	178.4	226.2	73.1
30-34	86.7	118.1	94.0	87.0	155.6	129.9	109.5	68.5
35-39	41.8	58.3	101.7	33.2	34.5	76.9	61.5	65.0
40-44	8.2	16.0	19.8	9.6	20.1	33.3	9.0	15.2
45-49	0.4	0.0	0.0	0.0	0.0	9.8	0.0	0.0
TFR	1.52	1.25	1.49	2.42	3.71	3.27	2.16	1.34

Own elaboration based on ONS LS, September 2010.

5.6.3 Conclusions

This section has outlined the percentage of births by selected countries of birth from the LS, compared to the ONS FM1 volume and then used the LS to calculate age-specific fertility rates for selected countries of birth. Using the all consistent cases category gives the most representative percentages from the LS when compared against the FM1 figures. However, in the latter part of the 1991-2001 period it is necessary to use the 2001 variable, as the 1991 variable does not give such suitable figures. This is likely to be because of changes in the LS composition in the latter part of the decade with fertility among entrants to the LS in the 1990s who were not at the 1991 census and so for whom the country of birth details are unknown. The best reading of the tables comes from using the 1991-1996 values from the 1991 census country of birth variables and the 1996-2001 period with the 2001 country of birth.

Type 1 consistent cases (resident throughout the 1990s) do not give very comparable percentages. The percentage of births for UK born women are much higher than the official figures while for other countries the percentages are much lower, especially for

India, Pakistan, East Africa and the Other European Union category. To an extent this must be a reflection of the characteristics of the sample – the percentage of LS members who are a Type 1 case is probably highest among the UK born group.

Age-specific fertility rates were calculated for the same mother's country of birth as in the percentage calculations. As explained, there is not an official statistics comparator for England and Wales. For India, Pakistan and Bangladesh there are dramatic age-specific trends which can be seen. India and Bangladesh both show early peaks in their fertility rates – the highest ASFR is in the 20-24 age group for these countries. For women from these countries the rates are very high in the key childbearing ages (20-39) and then high even in the 40-44 age group in the case of Bangladesh. For the Other European Union there are high rates of teenage fertility in some years and a late peak in rates.

5.7 Conclusions

This chapter has used the residence trajectories developed in the last chapter to calculate fertility rates and identify where deviations in these from the official rates are due to denominator and numerator differences. Research using the LS has preferred to take a cohort perspective to analysis. This, however, requires an awareness of the way in which the LS is comparable to national statistics, and it is therefore accurate from which to make inferences for the whole population. Often, the large size of the dataset is discussed but without note of information on data quality and representativeness. Analysis shown here has taken the ONS LS method for assessing 'sampling fractions' and 'linkage rates' and applied this to the two broad types of cases identified in Chapter 4. This has allowed a comparison of the impact of including selected inconsistent cases in the sample for analysis.

Crucially, in this chapter the fertility rates calculated have been arrived at through understanding the full exposure to risk for LS members between 1991 and 2001. Fertility rates using all the consistent cases are comparable with the official figures.

Although in some age groups the use of the Type 1 (continuously resident) types gives ASFRs which are closer to the official figures throughout the childbearing years, the all consistent cases have a better fit. Figure 5.1 showed that there is a high level of comparability on average throughout the 1991-2001 period. In comparing the denominator with the ONS mid-year estimates it has become apparent that later in the 1990s the use of the all consistent cases category gives a better fit with the target sampling rate. This is probably related to the attrition of LS members in younger age groups, as was discussed in Chapter 4. Births for all consistent cases in the LS show a better average fit with the 1.09% sampling target than the continuously resident type.

When selected inconsistent cases were included in the analysis the ASFRs dropped slightly and the estimated TFR was lower than the official figure from the ONS. This suggests that the use of the all consistent cases in analysis is preferable because there is a better match between the denominator and numerator. While sampling fractions for women are higher when the inconsistent cases are included, there is still a lower rate for the 25-34 year old age group. Sampling fractions for births using consistent and inconsistent cases are also higher when compared to the consistent cases on their own.

Births by selected countries of birth were compared to percentages in the FM1 volume and ASFRs calculated. Percentages of births by selected countries of birth are comparable with the ONS FM1 volume but use of the vital statistics country of birth variable for the mother would eliminate the decline in percentages from 1991-2001 when the 1991 census variable is used and then the increase from 1991-2001 when the 2001 census variable is used. Type 1 consistent cases (resident throughout the 1990s) do not give very comparable percentages. This shows that to understand the fertility among non-UK born women accurately it is necessary to include women who enter the LS after the 1991 census.

Age-specific fertility rates for the same countries as the percentages were calculated. India, Pakistan and Bangladesh show dramatic age-specific trends. India and

Bangladesh both show early peaks in their fertility rates – the highest ASFR is in the 20-24 age group. The Other European Union category shows a high rate of teenage fertility in some years and a late peak in rates.

Chapter 6

Assessing the suitability of the ONS LS for estimating the fertility of recent migrants – is there bias in entry to the LS among migrants?

Chapter abstract

This chapter is concerned with identifying if there is bias related to the entry of migrants into the Office for National Statistics (ONS) Longitudinal Study (LS) and the subsequent fertility of these new LS members. Three tests are applied to women in the ONS LS at the 2001 census. The first test identifies how well the ONS LS captures female migrants in intercensal years relative to the 2001 census, the second of these is concerned with identifying if there is a difference in the number of births around the time of the 2001 census, depending on the form of entry and the third, test in this chapter estimates the duration from registration with a GP and entry into the ONS LS to subsequent birth.

6.1 Introduction

As part of the over-arching objective to use the Office for National Statistics (ONS) Longitudinal Study (LS) to estimate fertility trends among new migrants to England and Wales, it is necessary to consider potential sources of bias in the dataset. Therefore, the aim of this specific piece of work is to identify and quantify any systematic bias in the data related to the form of entry which the LS member takes in the 1991-2001 period, to calculate the number of births to LS members depending on the form of entry taken and to calculate the duration from entry to birth among new female LS members.

The next section details terminology on migration and entry to the ONS LS; the process through which a migrant enters the ONS LS and is recorded in the dataset is not simple. Terminology to be used in the analysis in this chapter and the rest of the thesis will be explained in full and standardised to avoid confusion. Section 6.3 gives a full explanation of the reasons for testing the entry of LS members in relation to the census and their subsequent fertility. In section 6.4 the methodology for the tests which have been devised is explained in detail. Analysis of the results is in section 6.5, before the implications of the findings are considered in section 6.6.

6.2 The process for migrants entering the ONS LS and terminology to be used

This section details the purpose of this analysis and the background to the three tests which will be run to try and identify any possible bias in the LS related to new entrants and their subsequent fertility.

6.2.1 Distinguishing between entry to England and Wales and entry to the ONS LS

In using the ONS LS it is important to consider the terminology used to describe the movement of migrants into the LS from their arrival in England and Wales. The process of adding migrants into the LS from their arrival is easy to misunderstand.

This section explains the terms which will be used in subsequent chapters of this thesis and precisely what these terms mean. The differences in the terms are in many cases subtle, but very important in terms of the date at which they occur.

1. Date of migration to / entry to England and Wales
Term to be used – ***'Date of migration to England and Wales'.***
Measurement in the ONS LS – ***none (information from decennial census on if the LS member was overseas 12 months before).***

- The date of migration to England and Wales is the first date when the LS member arrives in England and Wales for the first time. This would be the date on which the person arrives in the country via whatever means of transport are used (e.g. date of first arrival in England and Wales – 8 July 1998).
- This is the true date of migration and, for this work, that of primary interest as we are concerned with the duration from the date of migration to England and Wales to birth for migrants.
- The actual date of arrival in England and Wales is not recorded by the ONS LS.
- From the use of the question at the 2001 census on the place of residence 12 months before it is possible to measure whether the LS member migrated from overseas within 12 months of the census.

2. Date of registration with an NHS GP
Term to be used – ***'Entry into the ONS LS'.***
Measurement in the ONS LS – ***recorded by the NHSCR; year, month and day of registration.***

- When a migrant to England and Wales with an LS date of birth registers with a GP for the first time the date on which the GP registration happens is recorded on the NHSCR (e.g. date of registration with a local GP in Southampton – 8 September 1999).
- There is no way of knowing what the duration is between the date of migration (1 above) and the date of registration with a GP, which is the date of entry to the ONS LS. This duration / exposure is unknown.

3. Date of tracing on the NHSCR (entry into correspondence between numerator and denominator

Term to be used – *'Date of trace'.*

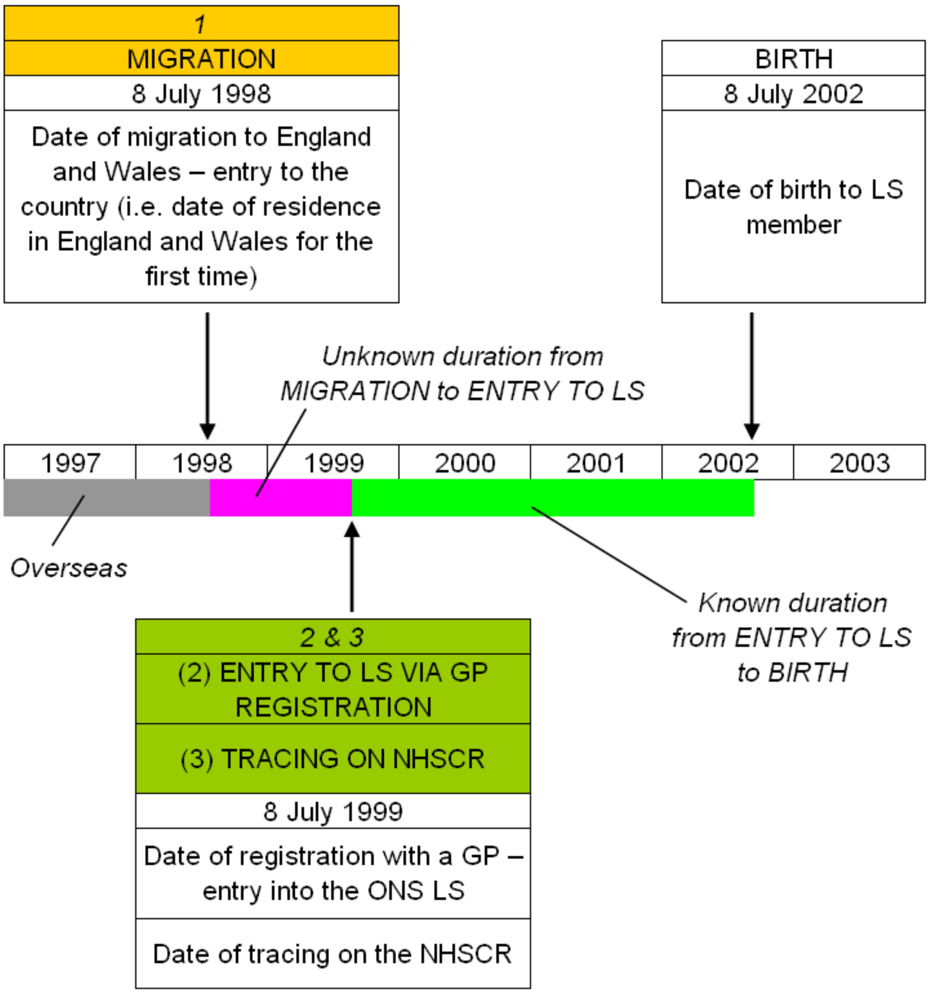
Measurement in the ONS LS – *recorded by the NHSCR.*

- Tracing is an activity which is ongoing and integral to the development of the LS. It refers to the flagging on the NHSCR of LS members who should have events data attached (i.e. births, deaths, embarkations, cancer registrations).
- Tracing can occur at the time of the census, when LS members' records are found in the NHSCR, and between census dates when an event occurs to an LS member who had not been traced until that point.
- Events can only be added to LS members after they have been traced. Therefore, the date when the LS member was traced is important.

While this thesis is concerned with the date of migration to England and Wales (1) in the ONS LS, this data is not available. The date of entry onto the NHSCR (2) can be used as a proxy or give an indication of the approximate date of arrival to England and Wales. This information must be used carefully with the trace on the NHSCR (3) to ensure that there is correspondence between the numerator and denominator. Persons who register a birth and give an LS date of birth will not become an LS member and that event will not be recorded.

Figure 6.1 shows diagrammatically the difference in the meanings of these terms. The first point in the figure, the date of migration to England and Wales, cannot be measured. This is on the extreme left of the diagram. In the ONS LS the entry to the LS is through registration with an NHS GP. This can be taken as a proxy for entry to England and Wales but this is assumed rather than known for sure. There is the potential for an unknown duration from entry to the country to the registration with a GP and entry into the ONS LS – this is denoted by the pink line between 1998 and 1999. At the date of entry onto the NHSCR tracing on the NHSCR is also attempted. On the diagram the duration from entry to the ONS LS to birth is shown in order to illustrate the known duration in the dataset.

Figure 6.1: Terminology in the migration process for migrants entering the ONS LS



Source: James Robards, August 2011.

Table 6.1 presents the terminology which will be used to refer to each of the events which have been described above. This should minimise potential confusion arising from the different entry stages in the ONS LS.

Table 6.1: Terminology on the entry of ONS LS members (migrants) used in this thesis

Event	Terminology used in this thesis	Notes
Migration to England and Wales	*'Date of migration to England and Wales'*	Not recorded by the ONS LS.
Registration with an NHS GP	*'Date of entry to the ONS LS'*	Recorded on the NHSCR and in the ONS LS.
Tracing on the NHSCR	*'Traced on the NHSCR'*	Recorded in the ONS LS at a broad decade level in the coding.

Source: James Robards, July 2011.

6.3 Background / rationale and research questions – why there is a need to test the entry of ONS LS members and their subsequent fertility

This section details the purpose of this analysis and the three tests which will be used to identify any possible bias in the LS related to new entrants and their subsequent fertility.

6.3.1 What is the purpose of this analysis?

Based on the events that an LS member may experience, it is possible that there is a differing likelihood of entry into the LS. Because the LS is composed of events data and is a non-consenting dataset it is possible that the greater the risk of the event, the greater the risk of entry into the study. There are two main ways in which there could be a bias manifest in the data as a result of the construction of the LS. These are explained below, along with a third question related to the duration from entry to the LS to subsequent fertility. This third question is concerned with the identification of an elevated level of first births related to the migration event and whether the exposure to risk of birth is a function of fertility.

As explained in section 6.2, because of the data used to construct the LS it is necessary to further understand the way in which the LS captures new migrants in the intercensal period with reference to the census. It is possible that there could be a bias in the data in terms of the entry of LS members, with those of childbearing ages who wish to access reproductive health services and women who intend or plan to give birth being more likely to enter the LS through registration with a GP. At each census there are women who are identified for the first time who did not register at a GP in the intercensal period but who become LS members as of the census. By comparing the number of women of each age group and their entry routes to the LS along with their births around the census and their fertility behaviour, the way in which this affects their entry to the dataset becomes apparent. Any bias in the selection of migrants into the LS should then become apparent.

6.3.2 Registration with a GP and entry on to the NHSCR – Research question: *What is the ratio of new entrants to the ONS LS in years prior to the 2001 census to the number of new entrants at the 2001 census?*

The first of the three potential sources of bias relates to the registration with a GP and entry onto the NHSCR. In Chapter 5 the number of LS members who were first identified at the 2001 census and entered into the LS was calculated. These LS members did not enter the dataset before the census. It is important to identify among female LS members how many cases there are relative to the number of female LS members who enter routinely via registration with a GP.

By comparing the number of LS members who enter the LS in the years preceding the census and the number of LS members who become LS members for the first time through addition to the LS at the 2001 census, it is possible to identify how good the LS is at capturing new members. The ratio of LS members entering in years preceding the 2001 census year to the numbers who are picked up at the census for the first time illustrates how good the LS is at capturing new members. Trends by age group can be assessed as all LS members have their year of birth recorded. The ratio calculated gives the ratio of LS members who enter in a manner which is expected and 'ideal' in that they register with a GP and enter in the expected, routine way (but not necessarily at the time of their entry to England and Wales) and those who enter at the census when they should have arrived in the LS at some point before.

6.3.3 Births to LS members around the 2001 census – Research question: *Are women who plan to give birth more likely to register with a GP and therefore enter the LS?*

A second unknown in using the LS is whether there is a bias in the type of LS member entering the dataset routinely – do women who are trying to become pregnant, or plan to give birth, enter at a higher rate than those who do not? Or is entry to the ONS LS a function of the intended fertility of migrants? Again, this is a testable scenario. The sample for this test is all LS members at the 2001 census who entered England and Wales in the preceding year (2000), in 2001 or entered at the 2001 census and traced on the NHSCR for the first time at the census. Using births in 2001 and births in the second half of 2001 (July-December 2001) the relative concentration of fertility between those who register and those who did not enter until the census can be calculated. This should show the degree to which the year of entry to the LS variable is accurate and how the births to LS members who enter the LS through the recorded entry date compare with those who enter at a census. The same analysis can be conducted for 2002 to see how the form of entry in 2001 impacts on the subsequent fertility of female LS members.

6.3.4 Duration to first birth among migrants – Research question: *What is the duration to first birth among new entrants / recent migrants to the LS?*

The last point of interest is close to the main research question; what is the duration to first birth among new entrants in the intercensal period? By calculating the months from the entry to the LS to the first recorded birth, it is possible to identify if there seems to be any rise in fertility after the migration event which would be consistent with findings in literature on this (Toulemon, 2004; Andersson, 2004). In addition to this, it is beneficial to estimate in months the duration as this gives an indication of whether there is a bias in the entry to the LS – do women who are new entrants to the LS enter (via a GP registration and thus entry to the LS) when they know that they are pregnant. Therefore, any increase in fertility which might be observed would be an 'artefact' of the way in which the registration process operates. Some migrant women may register with a GP only when they are aware that they are pregnant or have started trying to conceive a child. This would mean that using the NHSCR date of entry as a proxy for migration is unreliable.

An important note is that the data as recorded in the LS would not have the true birth parity. The first birth in England and Wales may not be the first ever birth to the LS member – the reproductive history of the LS member is not recorded as the data is composed of vital event information, rather than retrospective reporting. To account for the full fertility history of the LS member, the household information of the LS member at the 2001 census can be used; through using the relationship variables and ages it is then possible to reconstruct the fertility history of the LS members.

There is a dearth of published research on the duration from entry to England and Wales to registration with a GP. Evidence from Office for National Statistics (ONS, 2011) research using the Migrant Workers Scan (MWS) (a subset of National Insurance Registrations for non-UK citizens who have registered for a National Insurance number) shows that a third of migrant workers registered with the NHSCR within three months. There is some other research using the Migrant Worker Scan which has focussed on the feasibility of using the data for estimating migration (Sharfman et al. 2010), but this data is not part of the LS and cannot be matched in for this research. This same work by Sharfman et al. (2010) has identified that, in comparison with the NHSCR, the MWS shows a greater proportion of moves into London than the NHSCR. However, this is concerned with internal migration rather than international migration. In the latter part of Sharfman et al. (2010), the lag between UK arrival and National Insurance Registration is explored. It was found that 54% of National Insurance Registrations occurred within the first six months of arrival and 75% registered within 12 months of arrival in the UK. A greater proportion of non-married women registered within six months of arrival in the UK, compared with married women. For the 2002/03-2005/06 period the lag in arrival to registration seems to have reduced to less than six months. The research suggests that this could be related to accession eight (‘A8’) migrants. Although this research by the ONS is concerned with the more recent period than that around the 2001 census (which is the period of interest here), the work gives an indication of migration with reference to the NHSCR data.

The next section outlines the method for each of the tests.

6.4 Method

This section explains the methodology and selection of LS members to answer each of the questions outlined.

6.4.1 Test one - Registration with a GP and entry on to the NHSCR

What is the ratio of new entrants to the LS in years prior to the 2001 census to the number of new entrants at the 2001 census?

The ONS LS collects new members through persons being born on an LS date, through entering England and Wales and registering with a GP with an LS birth date and through being resident at a census and giving an LS birth date. For the purposes of this work the primary interest is in those persons who move to England and Wales and register with a GP giving an LS birth date. However, the LS members who enter at the census and are therefore not collected from GP registration are of interest relative to the numbers entering via registration with a GP. The census, in addition to collecting a wide range of socio-economic information on LS members, can be viewed as a 'mopping up' exercise for collecting LS members who have not entered via routine registration with a GP in the intercensal period.

The number of new LS members entering in the years prior to the census, relative to the number who enter at the census, will be calculated. Through dividing the number of LS members who enter in a year prior to the census (the years 1996 – 2000 will be used) by the number of LS members who enter at the census, a ratio of entries to non-entries can be given. New entrants at the 2001 census will be those LS members who have not entered the dataset before at any point, have not been resident at a past census and were traced on the NHSCR as part of the 2001 census – NHSCR link. The analysis will be run by single years of age as of the 2001 census; this will show any age group trends.

At the 2001 census the tracing of LS members on the NHSCR with a date of birth which does not match that on the NHSCR was possible. Where an LS member entered with a date of birth which, according to the census form was an LS date but on the NHSCR the person does not have an LS birth date, a 'flag' has been provided in the dataset. This variable, 'DOBDISC0' is used in this analysis and these members are excluded from the denominator of LS members entering at the 2001 census, the rationale being that those LS members have an added level of inconsistency and mismatch which adds in another level of complexity / inconsistency.

6.4.2 Test two – Births to LS members around the 2001 census

Are women who plan to give birth more likely to register with a GP and therefore enter the LS?

Similar to the first question on entries to the LS, the methodology to answer this question must work around the 2001 census and entries to the LS. Therefore, births to LS members entering in 2000 and 2001 through registration with a GP and LS members entering through being identified at the 2001 census will be used. For each type of entry the outcome will be whether or not the LS member gave birth in 2001 or 2002. As a result of this analysis the percentage of LS members giving birth, compared to those who do not give birth, will be known. This will go some way to show whether the form of entry which the LS member takes has a link with the fertility observed. It is possible that the LS members entering through registration with a GP will have a higher likelihood of giving birth than those persons entered through being identified at the census. The identification of entry conditional on wishing to give birth will therefore be identified.

6.4.3 Test three – Duration to first birth among migrants

What is the duration to first birth among new entrants / recent migrants to the LS?

This test first calculates the time of arrival in months by converting the year to months and adding in the month of arrival of the LS member. The month of arrival is then added in to give a total number of months. From this, the same calculations are used to establish the year of birth in months, with the month of birth being added to this to give the time of birth in months. The birth months are then deducted from the entry months. This gives the duration from entry to the LS to subsequent first birth. The number of first births by month from year of entry will be given.

It will be interesting to look at the results of this analysis in relation to the literature as Andersson (2004) identified that, in Sweden (using registry data), most of the births to migrants were within the first 12 months of migration and it is identified that many of the children born in Sweden were actually conceived before the registered immigration in Sweden. Migration and marriage are thought to be interrelated and it is suggested that marriage migration may be related to the short duration from migration to birth among migrants.

6.5 Results

6.5.1 Results from test one - Registration with a GP and entry on to the NHSCR

What is the ratio of new entrants to the LS in years prior to the 2001 census to the number of new entrants at the 2001 census?

Figure 6.2 shows the ratio of women entering the LS prior to 2001 and recorded as being traced on the NHSCR for the first time between the 1996 and 2001 census dates, relative to those LS members who enter at the 2001 census and are traced on the NHSCR for the first time at the census. The thick black line running through the series shows the average for the years 1996-2000. Where the ratio is below 1 then more people were entered into the LS via identification and tracing at the 2001 census than registered in one of the years before the census. Clear from the graph is that in the key childbearing years more LS members entered the LS through registering with a GP than appearing at the census and entering the dataset. In the 18-28 age group (age as of the 2001 census) there were more women entering the LS through registration with a GP than entering through being at the census with an LS date of birth. Among ages outside of the 18-28 age group the ratios of entries are low, showing that more entered at the census. Female LS members aged over 38 at 2001 had much lower entry ratios – figures below 0.5.

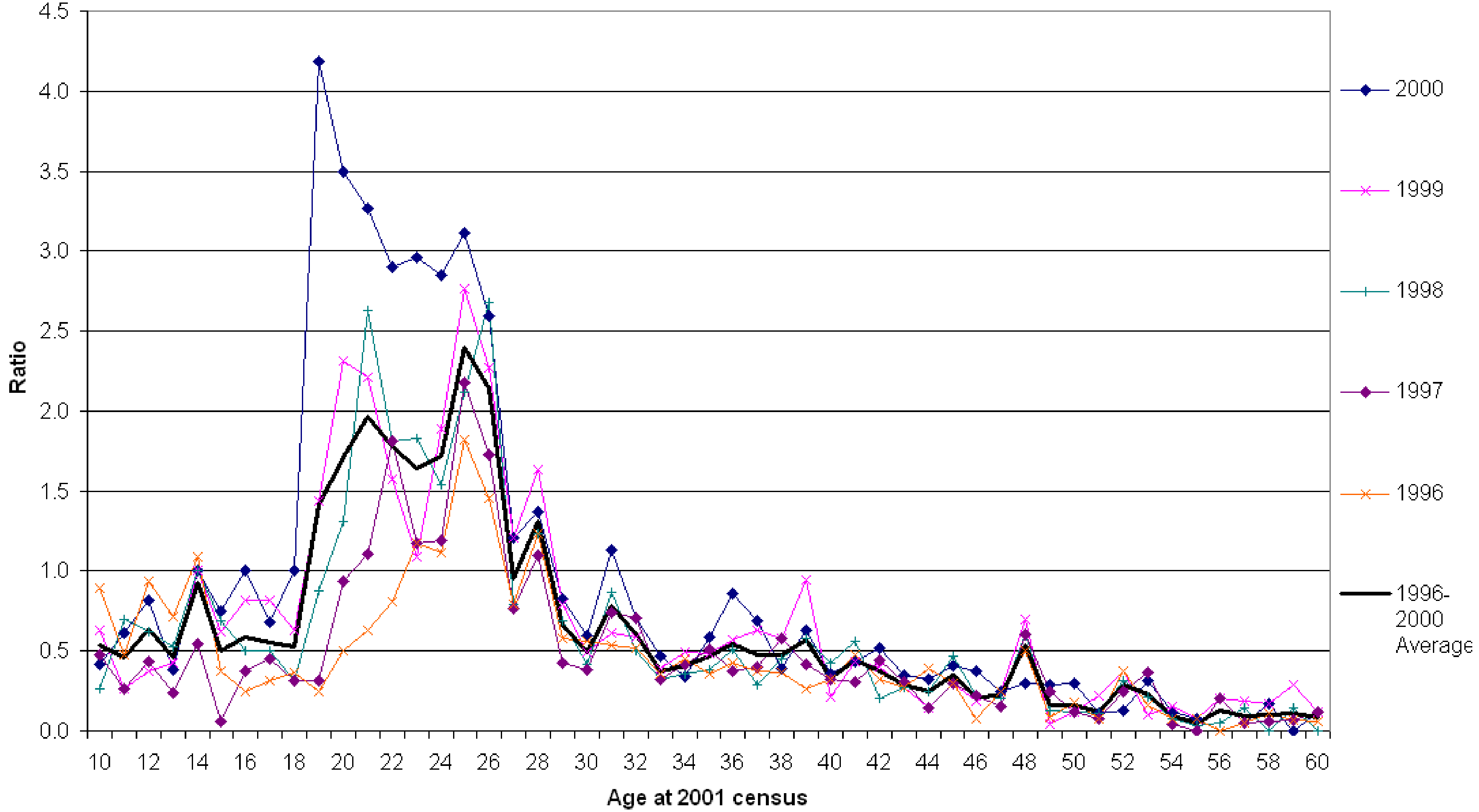
For the single years in the figure there is certainly a trend in the entry to the LS. There are high ratios of entry for LS members around the age of 18. For each of the years there is an increasing gradient in the ratio from age 17 to age 18 / 19, which coincides with possible demand for reproductive health services. In the series of years used, the year 2000 stands apart from the others with a very high ratio for LS members aged 18 and higher ratios for female LS members aged 20-25 than the other years in the series. This may be related to higher rates of migration around this time. When the years 1998 and 1999 are examined closely there are higher ratios relative to 1996 and 1997. This could be associated with the higher rates of migration which were observed in the late 1990s.

The number of entrants in the 18-28 age range, relative to those who arrive at the census, shows that there is a difference for this age group relative to the others in this series. It seems that women in the key reproductive age groups are more likely to register with a GP and enter the LS than women in older and younger age groups in the analysis. This finding links back to the work in Chapter 6 on the different forms of residence in the 1991-2001 period – for women who were a continuously resident consistent (Type 1) case, the lowest proportions were in the same age groups. With the inclusion of all consistent cases the number of women in these age groups increased greatly. This shows the importance of including new LS members in these age groups.

In assessing the importance of this main finding about the age of new LS entrants and the ratios calculated, it is important to remember that the numbers of LS members who arrive at the 2001 census are high, relative to the single years, because persons at the 2001 census for the first time and not traced on the NHSCR may have arrived at any time in the preceding decade.

Similar analysis has used the NHSCR as part of the Improving Migration Statistics Programme (ONS, 2012). This work was completed by the Office for National Statistics Methodology Division Demographic Centre and Southampton Statistical Science Research Institute (S3RI).

Figure 6.2: Ratio of women entering the LS in a year prior to the 2001 census relative to those who are traced for the first time at the 2001 census and have no date of entry prior to the census – by age at 2001 census



Own elaboration based on ONS LS, February 2011. (LS members who report an LS date of birth at the 2001 census and enter the LS but do not have this same date on the NHSCR excluded)

6.5.2 Results from test two – Births to LS members around the 2001 census

Are women who plan to give birth more likely to register with a GP and therefore enter the LS?

The second of the tests which were outlined in the methodology section seeks to understand differences in births between women who enter the LS on a date prior to the census and those who enter at the census. This can be understood using the date of entry among LS members and also using the 'trace' variable which indicates the date when an LS member was first traced on the NHSCR. Through using a carefully defined criteria, LS members can be selected and the number of births to LS members recorded.

Analysis here is divided into LS members who enter in the year 2000, just before the 2001 census, and those members who enter in 2001. It is important to remember that this analysis is for an LS member who gave birth in comparison with a member who did not give birth, and not for first births or the total number of births. For each table the percentages of LS members which have a date of birth discrepancy are also shown (these are in grey as they are of less interest).

Entry in 2000

Table 6.2 shows the number of women who entered the LS in 2000 and those who entered at the 2001 census (and were added through the post-census LS-NHSCR linkage exercise during which they were traced). This seems to show that those who entered in 2000 have a higher rate of fertility in May-December 2001 than the LS members entering at the 2001 census. However, the duration to a birth is important to consider and therefore the comparability between the two groups is problematic. This analysis does not control for the time since entry to the ONS LS. Those who enter in 2001 were traced for the first time at the census (April 2001), while those who entered in 2000 have been resident for a longer period of time and are more likely to have given birth (see next section).

Table 6.2: Percentage of births in 2001 to LS members entering in 2000 and at the 2001 census

	BIRTH MAY-DECEMBER 2001	NO BIRTH MAY-DECEMBER 2001
RECORDED ENTRY IN 2000- dobdisc 0	7.1	92.9
ENTRY AT 2001 CENSUS- dobdisc 0	3.0	97.0
RECORDED ENTRY IN 2000- dobdisc 1	14.7	85.3
ENTRY AT 2001 CENSUS- dobdisc 1	0.2	99.8

Own elaboration based on ONS LS, February 2011. (Dobdic = 1 – date of birth on census form does not match that on NHSCR).

Table 6.3 uses the July-December part of 2001 to make the same comparison. The results show again that the LS members entering at the 2001 census were less likely to have had a birth in the latter part of 2001, relative to those LS members entering in 2000. As with Table 6.2, it is likely that the results are affected by the lack of comparability between the two entry dates.

Table 6.3: Percentage of births in the second half of 2001 to LS members entering in 2000 and at the 2001 census

	BIRTH IN SECOND HALF OF 2001	NO BIRTH IN SECOND HALF OF 2001
RECORDED ENTRY IN 2000- dobdisc 0	6.4	93.6
ENTRY AT 2001 CENSUS- dobdisc 0	2.6	97.4
RECORDED ENTRY IN 2000- dobdisc 1	11.8	88.2
ENTRY AT 2001 CENSUS- dobdisc 1	0.2	99.8

Own elaboration based on ONS LS, February 2011. (Dobdic = 1 – date of birth on census form does not match that on NHSCR).

Entry in 2001

Tables 6.4-6.6 show entries to the LS in the year 2001, the form which these took and the corresponding number of women who had a birth, compared to women who did not give birth. Most interesting from the tables for this test is Table 6.5 which shows births in 2002 based on the form of entry in 2001. This is probably the most useful of the tables in this series and shows that, consistent with the others in this section, members entering through the census show a lower likelihood of giving birth, relative to those who enter at another point in the same year. However, there is less of a difference in this case than for the other tables. In total, 5.9% of LS members entering in 2001 through a GP registration gave birth in 2002, compared to 3.7% of those who entered at the census. Using 2002 is preferable because this allows both the forms of entry in 2001 to be used and gives LS members a similar exposure or duration of residence in England and Wales.

To extend this analysis further it would be possible to restrict this analysis to a smaller sub-set of LS members entering in 2001 – perhaps using only those LS members who enter around the same time of the census but not entering at the census itself. The difference between these two types of LS member would be interesting to look at.

Table 6.4: Percentage of births in 2001 to LS members entering in 2001 and at the 2001 census

	BIRTH MAY-DECEMBER 2001	NO BIRTH MAY-DECEMBER 2001
RECORDED ENTRY IN 2001- dobdisc 0	6.7	93.3
ENTRY AT 2001 CENSUS- dobdisc 0	3.0	97.0
RECORDED ENTRY IN 2001- dobdisc 1	0.0	100.0
ENTRY AT 2001 CENSUS- dobdisc 1	0.2	99.8

Own elaboration based on ONS LS, February 2011.

Table 6.5: Percentage of births in the second half of 2001 to LS members entering in 2001 and at the 2001 census

	BIRTH IN SECOND HALF OF 2001	NO BIRTH IN SECOND HALF OF 2001
RECORDED ENTRY IN 2001- dobdisc 0	6.1	93.9
ENTRY AT 2001 CENSUS- dobdisc 0	2.6	97.4
RECORDED ENTRY IN 2001- dobdisc 1	14.3	85.7
ENTRY AT 2001 CENSUS- dobdisc 1	0.2	99.8

Own elaboration based on ONS LS, February 2011.

Table 6.6: Percentage of births in 2002 to LS members entering in 2001 and at the 2001 census

	BIRTH IN 2002	NO BIRTH IN 2002
RECORDED ENTRY IN 2001- dobdisc 0	5.9	94.1
ENTRY AT 2001 CENSUS- dobdisc 0	3.7	96.3
RECORDED ENTRY IN 2001- dobdisc 1	15.4	84.6
ENTRY AT 2001 CENSUS- dobdisc 1	0.1	99.9

Own elaboration based on ONS LS, February 2011.

In assessing the results of this test it is important to recall that in the findings from the first test it was identified that the characteristics of women entering through registering with a GP are different to those who enter at the census. Women entering in the intercensal period are more likely to be in the key childbearing ages, compared with LS members entering at a census. This is likely to have a bearing on the outcome of the test.

6.5.3 Results from test three – Duration to first birth among migrants

What is the duration to first birth among new entrants / recent migrants to the LS?

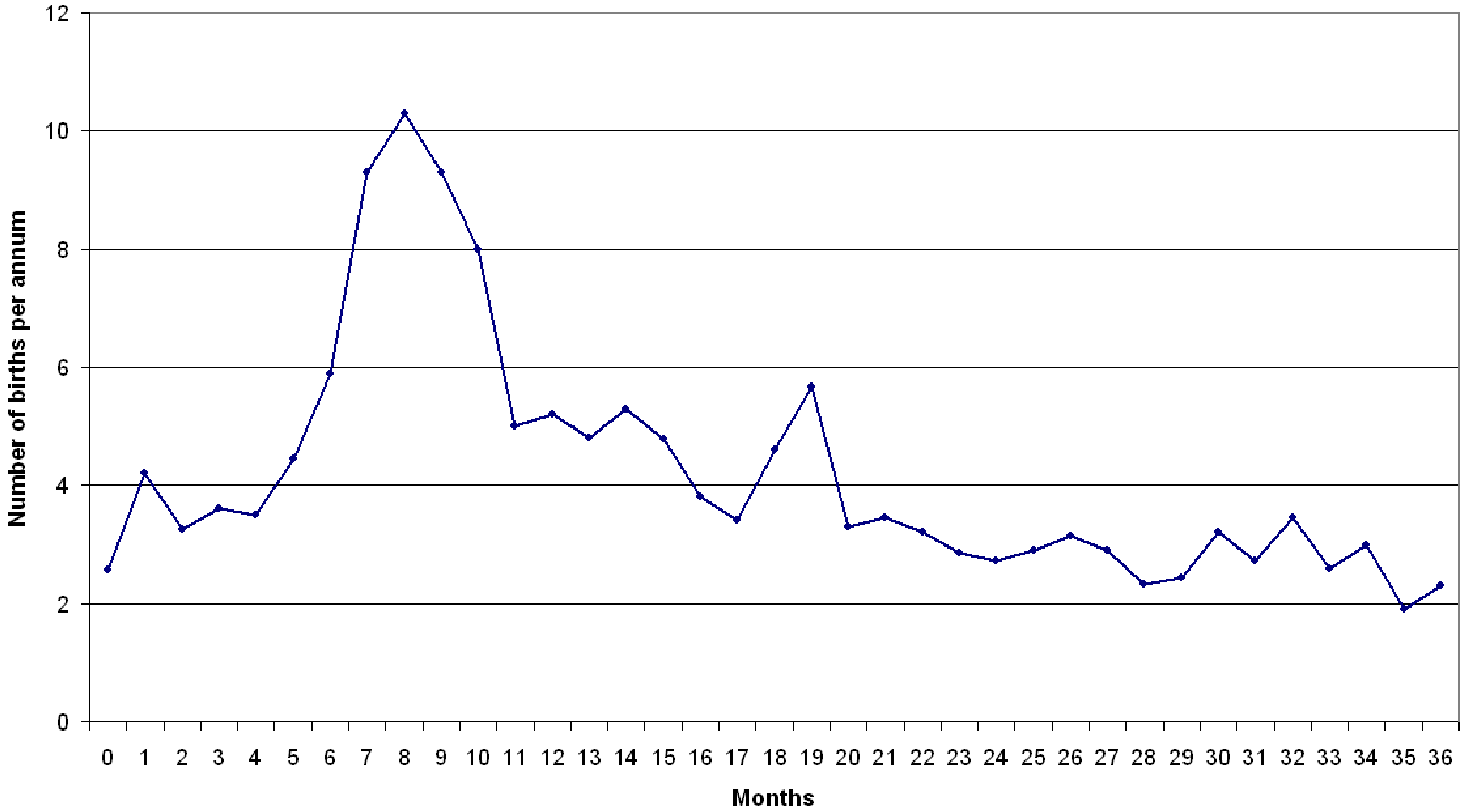
This test calculates the duration in months from the entry of the LS member to the study to the first birth. Analysis has also been run for the age of the LS member at birth, to give an indication of the groups which are contributing most to fertility. Figures 6.3 and 6.4 show the trends for 1991-2000 and 2001-2006 respectively. This test uses the year of entry for LS members who remain continuously resident from their entry into the study until the 2001 census, or LS members from 2001 who have not shown an embarkation.

Figure 6.3 shows the trend in the duration to first birth among new LS entrants who remain continuously resident in the 1991-2000 period. When the same analysis is run for a less restricted sample of LS members where any form of residence can be taken in the period the trend is similar (i.e. those members of a consistent or inconsistent type, who migrate again or disappear through attrition). The figure shows that there is a peak in first births to new entrants 8 months after entry to the LS. Following this, there is a steep decline in the number of first births to the 11 months age group. A second rise in first births is observed around the 18 months period, after which there is a gradual decline in the numbers to the end of the 36 month time period selected here.

The same analysis for the 2001-2006 period shows that there is a stronger peak in the first birth numbers, again in the eighth month after entry to the LS. This is a stronger trend and there are more births in this period of analysis. In this case there is no subsequent increase in the number of first births again with the numbers from the 17 month point averaging around 10 per annum for the rest of the period.

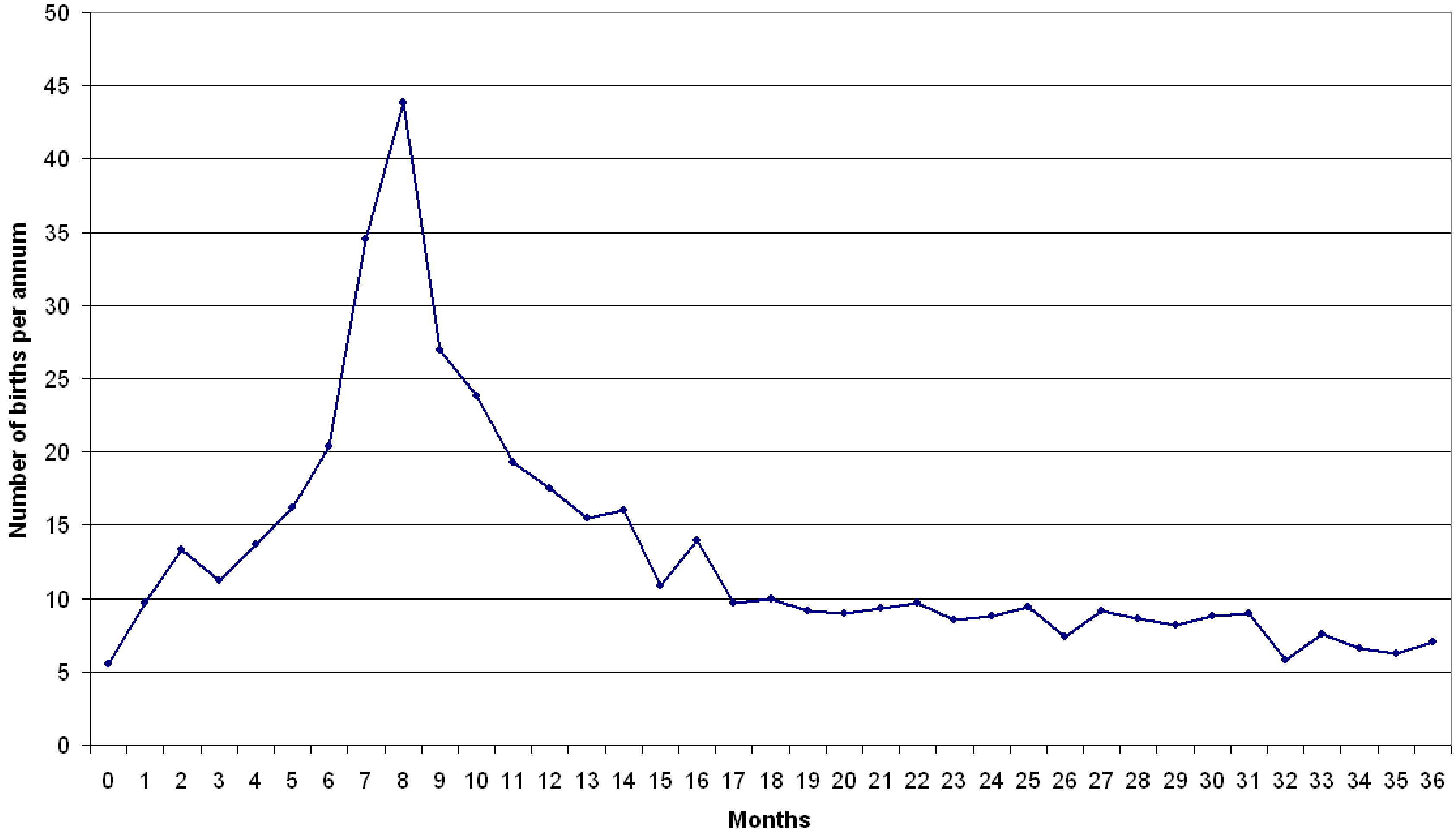
The results from this test show that there is a particularly high number of births in the eighth and ninth month after entry to the LS. This suggests that there could be a bias in the entry of LS members; women who know that they have conceived may register with a GP at that time when they have not registered up until that point. This means that the duration from the migration event to the registration with a GP is unknown and not shown in this analysis. Equally, it is possible that the fertility is associated with the migration event in some way. It could be the case that the fertility is related to migration because of marriage and then subsequent fertility, or that the migration and fertility are related to some kind of geographic preference over where the child is born.

Figure 6.3: Average number of first births per month after entry to the ONS LS for the period 1991-2000



Own elaboration based on ONS LS, February 2011.

Figure 6.4: Average number of first births per month after entry to the ONS LS for the period 2001-2006



Own elaboration based on ONS LS, February 2011.

Developing the above analysis, a version of this has been run with grouped ages and months since entry to the LS. In order for the outputs from this analysis to be cleared by the ONS (because of small numbers), grouping the months since entry to the LS has been necessary.

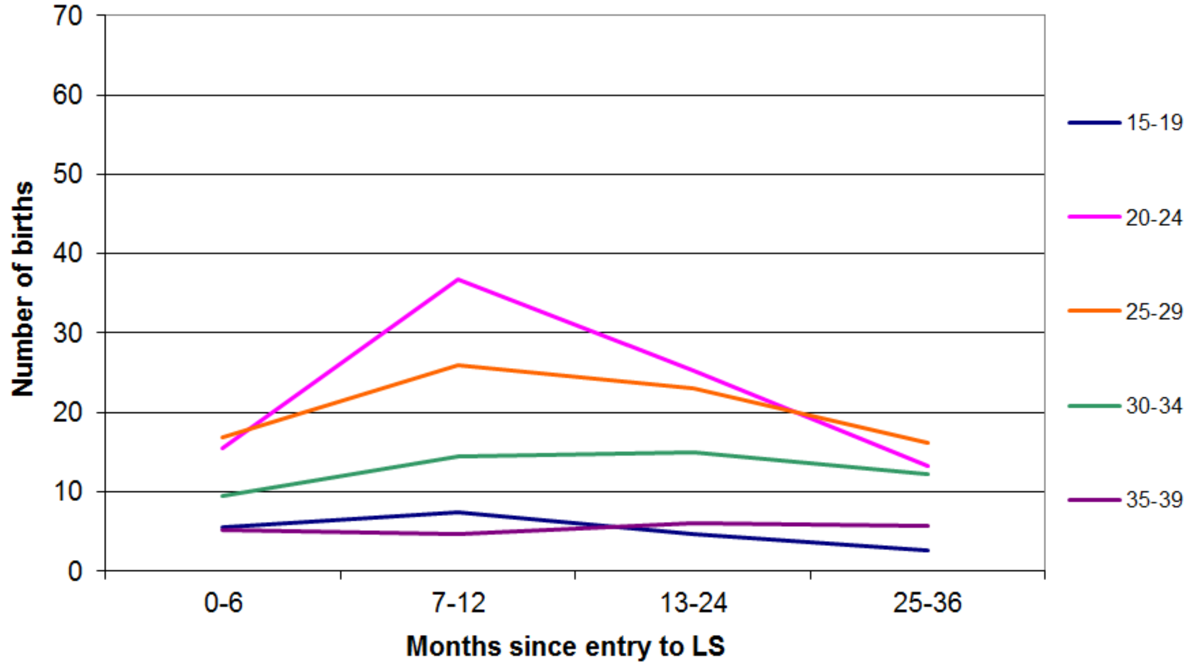
Figure 6.5 shows the data for the 1991-2000 period, while Figure 6.6 shows the same analysis for the 2001-2006 period. Both of these show that the highest number of births is in the 7-12 month period, which is consistent with the finding above.

The 1991-2000 period shows a picture where most of the births in this period are from the 20-24 and 25-29 age group. The 30-34 age group contributes fewer births. There is an increase in the births to all age groups, except the 15-19 age group from the 0-6 to the 7-12 month groups, although the 30-34 age group also shows another increase to the 13-24 months group.

In Figure 6.6 the trends are, as already explained, generally quite similar. However, the 25-29 age group show a similar rise to the 20-24 age group in the 0-6 to 7-12 months group and then also has a higher profile of fertility thereafter which is relatively higher than the corresponding groups in the 1991-2000 period. The 30-34 age group has a larger increase from the 0-6 to the 7-12 month group in this figure and this is also true for the 15-19 age group. In the 35-39 age group there is a slower rise to the 13-24 month category and no dip in the 7-12 month group, as in Figure 6.5.

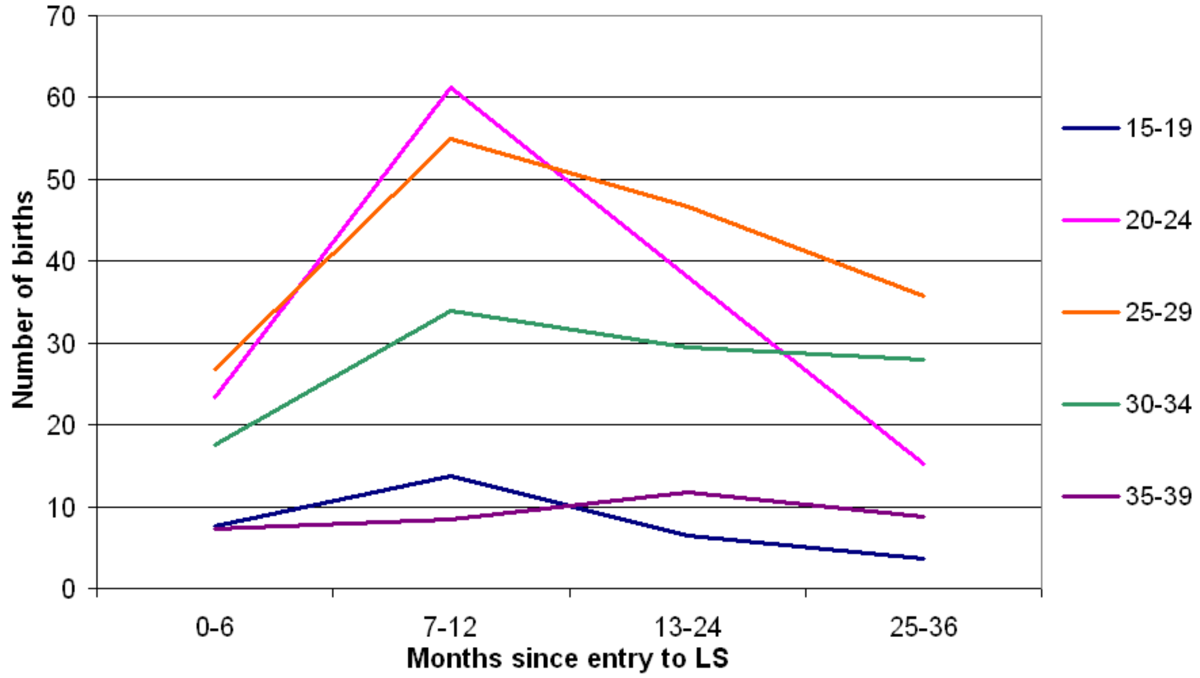
In some demographic research on migration and fertility, births in the first 9 months of residence have been censored. However, this would exclude the main finding of this analysis on the number of births in relation to the migration date.

Figure 6.5: Average number of first births per (grouped) month after entry to the ONS LS for the period 1991-2000 by age group



Own elaboration based on ONS LS, February 2011.

Figure 6.6: Average number of first births per (grouped) month after entry to the ONS LS for the period 2001-2005 by age group



Own elaboration based on ONS LS, February 2011.

6.6 Conclusions and implications

The purpose of this work has been to identify
any systematic trends in the registration of LS members with a GP, and whether their fertility intentions manifest themselves in the dataset. It is crucial to remember that the LS is composed of administrative data with each of the datasets combined to produce the LS collected for a different purpose. Key to the LS is the NHSCR which is where the matching, linking and tracing of LS members takes place. Because of the different meanings of the relatively similar terminology used to describe the LS, the first section of this chapter explained and standardised the terminology. The standard terms will be used in the remainder of this thesis.

Results from the first test on the numbers of LS members entering in the five years before the 2001 census, relative to the numbers entering in the census year, show that the LS mainly collects women in the key childbearing ages through routine registrations with a GP. For the younger and older ages there are lower ratios and therefore higher numbers entering at the 2001 census. This finding is consistent with trends in migration more generally – it is a process where most migrants are in the key childbearing age groups. The analysis in this section relates back to that in Chapter 5 on the forms of residence which LS members have taken between the census dates and the characteristics of the new female entrants in the intercensal period. It seems there is an association between GP registration and wanting to access reproductive health services.

The second test on the form of entry to the LS and subsequent births showed that there are higher rates of fertility for those LS members who registered with a GP and entered the LS outside of the 2001 census. However, this finding links back to the form of entry that the LS member takes and the findings from test one. It is known that most of the new entrants who entered through residence at the 2001 census were under age 18 or over the age of 28. Therefore, it would be expected that the results from this test show a lower level of fertility for entries through the census, relative to those LS members who enter in a routine way through registration with a GP.

Test three sought to identify if there is a higher level of fertility among new entrants to the LS. The analysis showed that among new entrants to the LS the greatest number of births to LS members come in the eighth month after the recorded entry to England and Wales, as sourced from the NHSCR. This suggests that the registration with the GP might be taking place at the time of conception, that the registration may take place once the LS member knows that they are pregnant (some unknown time after the date of arrival in England and Wales), or that the registration takes place on arrival in England and Wales and that this is related to some form of family reunification or post marital migration to England and Wales. It is very hard to identify further which of these may be correct. The main point is that the association between GP registration and giving birth suggests a selection effect where those women who give birth register around conception / pregnancy. Exposure to risk of birth among migrants in the LS seems to be related to registration with a GP, which is the main way in which migration into the LS is recorded. Indeed, the date of registration with a GP is called the date of migration to England and Wales, although it is not actually measuring this.

In order to develop the finding from test three that there is a high level of fertility within the first nine months of residence as recorded on the NHSCR, there are various options. The first of these would be to work at some form of broad level in terms of the duration of residence in England and Wales to minimise the potential dominance in analysis of fertility which is related to the registration with a GP. For example, the first 24 or 36 months of residence could be used as, if there were an unknown exposure of three months in the period before registration with a GP and entry to the LS, then this would be minimised. However, this would to an extent undermine one key objective of this work: to look at the duration from entry to England and Wales to subsequent births.

A second option would be to compare the duration from entry to birth in the LS with another dataset. The Annual Population Survey from the ONS (ONS, n.d.g) has a question on the year of entry of the respondent and also includes their fertility history. Somehow it may be possible to use this information to corroborate the findings made here on the ONS LS. A clear drawback of this approach is the inclusion of another dataset in this analysis.

Another option would be to use the household information from the 2001 census to reconstruct the fertility histories for LS members. This may give an indication of pre-migratory fertility as at the 2001 census the country of birth of the children to LS members was asked. At the 2001 census, the location of the LS member 12 months prior to the census is asked. It might somehow be possible to use this information relative to the date of entry into the LS and residence at the 2001 census.

In the case of all tests it is possible that the qualitative change in migration to England and Wales (and therefore registration with a GP) is important. For example, it is known that since 2004 there has been a high rate of migration related to the new accessions to the European Union in that year. This would mean that the migrant type is different to that which was entering in the period around the 2001 census. In comparing migration and subsequent fertility, the country of origin and therefore the cultural reasons for the migration of the LS member may have an important linkage to the duration to birth. For example, fertility among women moving from Asia may be related to marriage and subsequent migration, as Andersson (2004) identified in Sweden. This in turn links back to the likely duration from arrival in England and Wales to subsequent fertility and the possible unknown exposure (pre-GP registration), which is of concern.

Overall, the tests here have identified that there are notable trends in the LS for looking at the fertility of new LS entrants / recent migrants. Entry to the LS in non-census years is more likely for women in the key childbearing ages, entrants from outside of a census year are more likely to give birth and there is a trend in the fertility of new LS entrants whereby most births to LS members occur within the first 12 months of residence and predominantly in the first 9 months, suggesting that the conception might be related to the registration and entry into the LS rather than related to post-migratory fertility behaviour.

One of the most important implications of the findings here is that analysis away from a census date will not provide a sample with which it is possible to be certain about the date of migration. This means that analysis of the period after the 2001 census should be limited to the years immediately around the census (one to two years) otherwise there is the risk of selecting samples of LS members who are not resident in years away from census dates. In addition to the registration and birth relationship identified in this chapter, in Chapter 4 it was identified that among foreign born women there was a higher risk of attrition. Based on the findings made here, the next chapter selects samples of migrants and a comparator group of non-migrants, using information from the 2001 census form about the usual place of residence 12 months before the census. The use of the migration indicator at the 2001 census should assist with the selection of a sample independently of the date of registration on the NHSCR. Migrant and appropriate comparison groups will be identified for use in modelling the risk of birth after the 2001 census.

Chapter 7

Identifying a sample of recent migrants entering the ONS LS before the 2001 census and estimating their fertility

Chapter abstract

This chapter develops findings from Chapter 6 on the form of entry for Office for National Statistics (ONS) female Longitudinal Study (LS) members and their births around the 2001 census. Given the association between registration with a GP and subsequent birth it was concluded that using information on the whereabouts of the LS member 12 months before the 2001 census, together with other information that identifies recent migrants offers an alternative criteria for sample selection. Therefore, the first section of this chapter covers the selection of LS members resident at the 2001 census. Fertility rates for all groups in the years immediately after the 2001 census are calculated. Findings from this chapter provide important information on the selection of a sample and comparison groups for estimating the fertility of recent migrants to England and Wales.

7.1 Introduction

This chapter explains the selection of a sample of female Office for National Statistics (ONS) Longitudinal Study (LS) members for fertility analysis who were recent migrants as of the 2001 census. The key aim in the initial part of the chapter is to devise criteria for the selection of LS members who were at the 2001 census for the first time and appropriate comparison groups, which includes those LS members entering at some point in the 1991-2001 period and also those LS members who were continually resident consistent cases in the 1991-2001 period. Given the complexities in the way in which the LS functions and the previous findings of a possible association between entry to the ONS LS / registration with a GP and subsequent fertility (see Chapter 6) and attrition among foreign born women (see Chapter 4), the criteria for selection of a sample is detailed.

Section 7.2 provides an explanation of the research questions, aims of this work and the approach adopted, while section 7.3 considers methods to select a sample of migrants at the 2001 census, the comparison groups and explains the precise approach taken. Subsequently, section 7.4 profiles the groups devised to give a detailed socio-economic background, before section 7.5 details deaths and immigration among these groups in the period since the 2001 census. In section 7.6 the fertility for each group around the 2001 census (before and after the census) is detailed and fertility rates provided.

The migrant group identified for further analysis is composed of LS members who were living overseas 12 months before the 2001 census and traced on the NHSCR between 1991 and 2001 or at the 2001 census for the first time. Three comparison groups are identified: continually resident consistent cases between 1991 and 2001; LS members who entered the LS between 1991 and April 2000 but were not living overseas 12 months before the 2001 census and LS members entering the LS between April 2000 and April 2001 (the date of the 2001 census). Descriptive statistics are calculated for the 2001 recent migrants group and comparison groups. The main objective of this research to identify if there is a trend in fertility among recent migrants (or entrants to the LS) whereby there is an elevated level of fertility in the years after the migration event which is associated with the migration event. For this reason, in the latter part of this chapter there is some detail on the births to these samples of LS members.

Tables 7.1 and 7.2 provide summaries of the groups to be selected and the data from the ONS LS which is used in the selection of each group. Table 7.1 illustrates the four groups which will be identified and the sources of data which can be used to estimate the date of migration to England and Wales. The columns on the right show when it is possible to know if the LS member was resident. The 'date of migration' is established through the use of the date of entry to the ONS LS with the NHSCR registration and the date of trace. Table 7.2 gives more information on the date of trace which is used in the sample selection.

Table 7.1: Characteristics of groups to be selected

Group	Date of migration (not from ONS LS)	Date of entry to NHSCR (from ONS LS)	Date of trace (from ONS LS)	Present at 1991 census (April)	Present in March 2000	Present in April 2000	Present at 2001 census (April)
Analytic	Within 12 months of the 2001 census	NA - not used.	1991-2001 *Or* At 2001 census	No	No	Yes	Yes
Comparator 1	Before 1991 census *Or* None	Before 1991 census *Or* At 1991 census	Before 1991 census *Or* At 1991 census	Yes	Yes	Yes	Yes
Comparator 2	1991-2001	1991-April 2000	1991-2001	No	Yes	Yes	Yes
Comparator 3	1991-2001	April 2000-April 2001 (2001 census)	1991-2001	No	No	Yes	Yes

Source: James Robards, August 2011.

Table 7.2: Entry to UK and NHSCR / LS

Group	Entered UK	Entered LS (NHSCR registration)	Traced in NHSCR
Analytic	2000-2001	2000-2001	1991-2001 (including at the 2001 census)
Comparator 1	<=1991 census	<=1991 census	<=1991 census
Comparator 2	1991-2000	1991-April 2000	1991-2001
Comparator 3	1991-2001	April 2000-April 2001	1991-2001

Source: James Robards, August 2011.

The next section gives a more detailed rationale and explanation of the approach which has been adopted and key points to consider in using the ONS LS. In relation to Tables 7.1 and 7.2, the way in which sample selection is approached will be made clear.

7.2 Rationale and approach

7.2.1 The context

As outlined earlier in this thesis (Chapter 3), the LS is unique in the way in which members are added into the dataset using information from the NHSCR. In Chapter 6 it was shown that there are systematic trends in the timing of entry to the LS among female migrants. Most importantly, births to new entrants seem to mainly come after 8 months from the date of registration with a GP, suggesting that in the LS registration with a GP could be associated with pregnancy. Because of this finding it would be unwise to use the date of entry on the NHSCR and subsequent birth dates as a way of estimating the duration from entry to birth with absolute certainty.

Therefore, this chapter identifies a sample for further analysis by using the 2001 census as a datum. The 2001 census can act as a datum because at that point in time it is possible to be certain about which LS members are resident (given the findings on attrition for LS members who were at the 1991 census and attrition for new entrants to the ONS LS in the 1991-2001 period). In addition to this, socio-economic information collected at the census is available for ONS LS members (there is none for post 2001 migrants who enter the LS) and this is important in estimating the subsequent fertility of the recent migrants to England and Wales. Focusing on the two years following the 2001 census therefore gives a sample where attrition should be minimal, where there is complete socio-economic information and will allow the identification of fertility trends for recent migrants into the LS relative to selected reference groups. This means that the findings of this research are more robust and the samples selected fully understood.

Given the finding in the last chapter, showing that there could be a relationship between entry to the ONS LS and the duration to subsequent births which was illustrated in the duration of 8-9 months from the entry to the LS to birth, the analysis here uses the migration indicator at the 2001 census to select a group of migrants. In doing this the analysis wishes to separate the NHSCR date of entry to the LS from the birth date and remove any association between the duration from entry to the ONS LS and subsequent births. By using the address a year before the census (as recorded on the 2001 census form) it is possible to identify those migrants who migrated to / entered England and Wales between April 2000 and April 2001 with reference to their past location. The date of entry to the ONS LS can also be used in the analysis to identify when new LS members registered with a GP for the first time. One of the most recent papers to have used the ONS LS and NHSCR information to look at migration and registration with a GP is that by Smallwood and Lynch (2010). This used the address 12 months ago question at the 2001 census to explore migration of LS members and trends in the matching of the address a year before the 2001 census, the address at the 2001 census and the NHSCR address.

7.2.2 Aims of this chapter / approach

There are three aims in this chapter.

1. Identify how many LS members have a date of migration to England and Wales for the first time around the 2001 census from overseas and identify suitable comparison groups.
2. Fully detail the background socio-economic characteristics of the sample in 1.
3. Estimate the births and fertility rates for the groups identified in 1.

These aims are integral to the next chapter which wishes to estimate the risk of birth to recent migrants into England and Wales using event history analysis. In order to do this it is crucial to be patently clear on the composition of the samples drawn from the ONS LS members resident at the 2001 census.

The next section discusses the key considerations in the selection of an appropriate sample for analysis.

7.3 Selecting a sample of migrants for analysis and suitable comparison groups

In this section there are two key aims – firstly to explain the functioning of the ONS LS with regard to the collection of data on recent migrants in 2001 and secondly to define migrant groups and control groups at the 2001 census.

7.3.1 Key considerations in the selection of a sample for analysis

There are several key considerations that need to be addressed in order for the correct samples to be selected. These are discussed below.

Tracing of LS members on the NHSCR

As previously explained (see Chapter 3), tracing means that events (and their date of occurrence) (i.e. births, cancer registrations, embarkations and re-entries) can be recorded for each LS member. At each census new ONS LS members are traced on the NHSCR and between each census migrants to England and Wales are traced on the NHSCR, when they are recorded as registering with a GP. Events occurring before tracing are not identified or included in the ONS LS (e.g. a birth overseas). The ONS LS contains a variable indicating when the LS member was traced at NHSCR: at a census, between censuses, or not at all. The values in Table 7.3 below refer to the date when the LS member was first traced at the NHSCR.

Table 7.3: Trace variable coding

Coding	Meaning
0	Untraced
1	Traced at 1971 Census
2	Traced between 1971 and 1981 Censuses
3	Traced at 1981 Census
4	Traced between 1981 and 1991 Censuses
5	Traced at 1991 Census
6	Traced between 1991 and 2001 Census
7	Traced at 2001 Census
8	Traced after 2001 Census

Source: CeLSIUS Data Dictionary, Accessed June 2011. http://www.celsius.lshtm.ac.uk/newDataDict/dddrill2k.php?varname=TRACE&sqlname=CORE1.

The trace variable is used to select only those female LS members who entered either immediately before the census or at the 2001 census, as part of the NHSCR processing and tracing stage. Any LS members not traced as of the 2001 census but traced after this would not be able to have birth information attached (although it is unlikely that they would give birth as they are not registered with a GP). If an LS member was resident at the census and then registered with an NHS GP after the 2001 census then the date of entry to the LS would be after the census date (and the trace would be '8 - Traced after 2001 Census'). It is for this reason that among the migrant groups the trace variable is being used in this analysis (all LS members have a trace value). If the LS member was a very recent migrant at the 2001 census then they may have registered after the 2001 census but before the tracing is completed on the NHSCR. In this scenario the trace code would be '8 - Traced after 2001 Census'. Therefore, for the first group which we wish to select (entries in the 12 months before or at the 2001 census) it is necessary to restrict the trace to include only those LS members who entered the LS either before the 2001 census (1991-2001 – trace 6) and at the 2001 census (trace 7). The trace coding which should be used in each case is outlined in Table 7.4.

Table 7.4: Trace variable coding to be used in the selection of migrant and comparison groups at the 2001 census

Sample	Trace value to be used
1. Entering in the 12 months before the 2001 census and overseas 12 months before the 2001 census.	Trace > 5 and Trace <8.
2. Consistent, continually resident 1991-2001.	Trace > 0 and Trace <=5.
3. Entries between 1991 and 2001 and not overseas 12 months before the 2001 census.	Trace = 6.

Source: James Robards, June 2011.

Date of birth discrepancies at the 2001 census

As mentioned briefly in Chapter 6, if at the 2001 census there was a different date of birth on the NHSCR compared with that on the census form, this was recorded and flagged as a separate variable. It is possible for a person on a census form who has incorrectly given an LS birth date to enter the LS if other information from the census form (name, sex, postcode, postcode one year ago and 'person type' (private/communal establishment)) matches that on the NHSCR. In previous LS-Census link exercises (1974, 1981, 1991) where there was a trace using a date of birth that does not match the NHSCR date of birth, there was no variable to give any information on whether the date of birth matched. In 2001, because of changes to the tracing process in the 1990s (computerisation in the early part of the decade) it was possible to include a variable which identifies where the LS member does not actually have an LS date of birth on the NHSCR at the 2001 census.

Where an LS member does not have an LS date of birth on the NHSCR but gave an LS date of birth on the census form, births to these LS members would not be recorded; for this reason it is important to use the variable which identifies where there is a mismatch in the date of birth and remove these individuals from the analysis. One way in which a discrepancy of this type can arise is through the completion of the census form by a non-LS member (who inadvertently reports on the census form an LS date of birth for one of the members of the household). Most entries to the ONS LS at the census (i.e. new LS members identified through giving an LS date on the census form) are through this form of discrepancy between the census form date of birth and the NHSCR date of birth. With regard to the selection of a group of migrants this could bias the sample (because subsequent birth events would not be recorded). This variable is only available for LS members at the 2001 census which makes it of further attractiveness for use in the selection of a sample from the 2001 census. Anyone who had a date of birth discrepancy at the 2001 census between the date on the NHSCR and on the 2001 census form will not be included in the analysis.

Date of entry to the ONS LS / registration with a GP

Because of findings that there is an association between registration with a GP and the time to the first birth (see Chapter 7), this work is concerned with using a duration to birth measure which is not influenced by the date of entry into the ONS LS. The date of entry onto the NHSCR will not be used in the selection of the migrant group as this may create a bias in the duration to first birth among LS members. However, as already detailed, it is necessary to use the trace variable and this is related to the registration with a GP.

7.3.2 Selecting migrant groups using the specified criteria

This section identifies a group at the 2001 census who migrated to England and Wales within the 12 months preceding the census and two comparator groups. Through the creation of a selected number of migrant groups the identification of a group who entered the LS before the 2001 census is possible. The subsequent fertility of these LS members can then be estimated. For this analysis the 2001 census is an appropriate datum as all LS members are accounted for at the census and new entrants not entered through another form of entry in the intercensal period would become apparent at this time. In the context of recent migratory trends (see Chapter 2) this is also the point where migration to England and Wales increased sharply. Through using LS members resident at the 2001 census, the full census and household non-member information for LS members can be utilised, including important variables such as country of birth of the LS member, marital status and parity.

Trajectories for migrant and non-migrant comparison groups at the 2001 census

This sub-section outlines the migrant and comparison groups which can be selected for analysis and uses a schematic plan (Figure 7.1) to illustrate the form of residence an LS member may have taken before the census.

- **Migrants entering the ONS LS in the 12 months before the 2001 census**

Group 1 – includes ONS LS members who indicated at the 2001 census that they were located overseas 12 months before (i.e. these persons entered the LS before the 2001 census). Entry to England and Wales must have been made between 28 April 2000 and 29 April 2001. There was no date of birth discrepancy at the 2001 census. These LS members entered the study through an NHSCR registration at some point in the 12 months before the 2001 census or entered at the 2001 census for the first time. There was no residence at a past census. These LS members were first traced on the NHSCR between the 1991 and 2001 census dates or at the 2001 census.

Therefore, tracing could have occurred before 2000 but this does mean that there is no fixed date of entry to the LS via the NHSCR and therefore the date of entry should be independent of any births. We take the information relating to the whereabouts of the LS member 12 months before the 2001 census as the priority – these persons said they lived overseas 12 months before the 2001 census.

- **Comparator groups – LS members who are continuously resident consistent cases (1991-2001) and migrants entering the LS (1991-2001)**

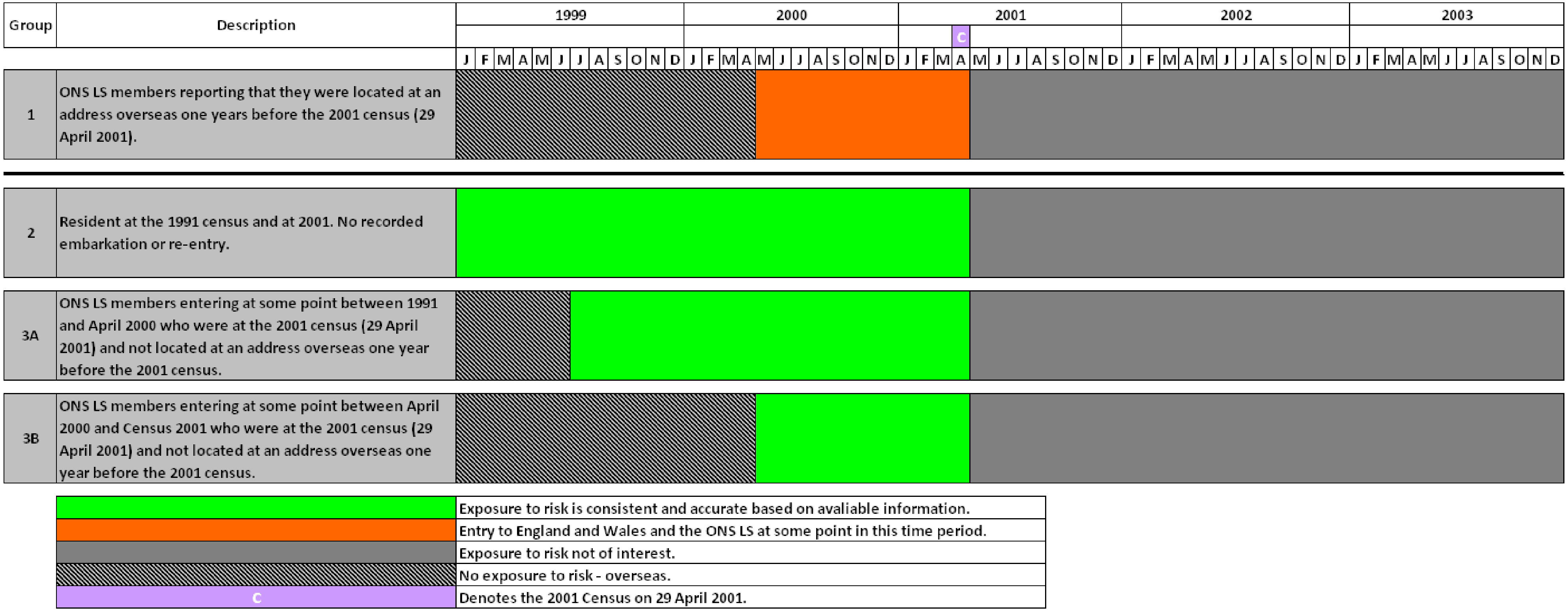
Three comparison groups can be identified. The first of these (Group 2) are LS members who are continually resident consistent cases (refer to Chapter 4 for an explanation of consistent cases) between the 1991 census and the 2001 census. The second of these (Group 3) are LS members who were living in the UK 12 months before the 2001 census but who had entered the LS via an NHSCR registration between the 1991 census and the 2001 census. As defined below, the third group can be sub-divided according to whether the date of entry to the ONS LS was pre- or post- April 2000. Along with the migration indicator, this tells us the most recent entrants to the ONS LS in the year before the 2001 census.

Group 2 – ONS LS members who were continually resident consistent cases between 1991 and 2001. There was no embarkation from or re-entry to the study and they were at the 2001 census.

Group 3A – ONS LS members who entered the ONS LS between the 1991 census and April 2000 (one year before the 2001 census) and were not located overseas 12 months before the 2001 census. They were traced on the NHSCR between the 1991 and 2001 census dates.

Group 3B – ONS LS members who entered the ONS LS between April 2000 and April 2001 (using the full date of the 2001 census) and were not located overseas 12 months before the 2001 census. They were traced on the NHSCR between the 1991 and 2001 census dates.

Figure 7.1: A schematic plan of residence patterns for migrants to England and Wales 12 months before the 2001 census and possible comparator groups



Source: James Robards, June 2011.

Table 7.5: Terminology on the entry of ONS LS members (migrants)

Sample	*Date of migration to England and Wales*	*Date of entry to the ONS LS*	*Traced on the NHSCR*
1. Entering in the 12 months before the 2001 census and overseas 12 months before the 2001 census.	- April 2000 – April 2001. - Reported as being overseas 12 months before 2001 census	- Between April 2000 and April 2001. - Or at the 2001 census.	- Traced in the 1991-pre-2001 census period. - Or traced at the 2001 census.
2. Consistent, continually resident 1991-2001.	- Before the 1991 census / not a migrant.	- Before the 1991 census. - May not be a migrant.	- Traced before or at the 1991 census.
3A. Entries between 1991 and April 2000 and not overseas 12 months before the 2001 census.	- Since the 1991 census as they were not resident at the 1991 census. - More than 12 months before the 2001 census.	- Entry between the 1991 census and April 2000.	- Traced between the 1991 census and the 2001 census.
3B. Entries between April 2000 and the 2001 census and not overseas 12 months before the 2001 census.	- Since the 1991 census as they were not resident at the 1991 census. - More than 12 months before the 2001 census.	- Entry between April 2000 and April 2001 (2001 census).	- Traced between the 1991 census and the 2001 census.

Source: James Robards, July 2011.

Post-2001 census residence patterns

Because the period of observation for these LS members will be the 24 months after the 2001 census, it is important to identify how many LS members leave in that period. In the latter part of this document the groups identified will be cross-tabulated with whether there was a death or embarkation in the years after the 2001 census. The next section presents the results for each group of LS members and their socio-economic characteristics.

7.4. Migrant numbers at the 2001 census and socio-economic background information

This section presents the number of LS members and the key covariate information for each of the groups identified in section 7.3. The age profiles of each of the groups, and the prevalence of students within the group, are key considerations. This information is important for the next steps in this research which will model the duration to first birth among a sample of recent migrants.

7.4.1 Overall numbers of LS members by group

Table 7.6 shows the numbers of LS members in each group, regardless of their age at the 2001 census. There are 875 female LS members who fall into Group 1. There are 169,421 female LS members in Group 2. Group 3 gives the number of LS members who entered between the 1991 census and the 2001 census and were not overseas one year before the census. This is split by the number of LS members entering between the 1991 census and April 2000 (12 months before the 2001 census) and those LS members who enter between April 2000 and the 2001 census (April 2001).For Group 3 there are a total of 6,579 LS members with 5,917 entering between the 1991 census and April 2000 and 662 entering between April 2000 and the 2001 census. The figure for 3B is very similar to the annual average for 1991-2000 for 3A.

Table 7.6: Numbers of ONS LS members per residence pattern

Group		Number
1	LS members at the 2001 census who gave an address a year before	875
2	LS members continually resident, consistent cases 1991-2001	169,421
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	5,917
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	662
Total		176,875

Own elaboration based on ONS LS, April 2011.

7.4.2 Numbers of LS members by trajectory and age at the 2001 census

Using age as of the 2001 census, it is possible to disaggregate the above categories, Tables 7.7 and 7.8 show the numbers and percentages of LS members in each group. The percentages are plotted in Figure 7.2. LS members who are continuously resident consistent cases have a flat age profile while the LS members who were overseas 12 months before the 2001 census are predominantly in the 20-24 years age group. Some 26% of this group are in the 20-24 years age group and 19% in the 25-29 years age group. LS members in the third type (entries between 1991 and 2001 but not overseas 12 months before the 2001 census) were slightly older, but showed a similar age profile to the new entries before the 2001 census. For Group 3B the age profile is very similar to that for the LS members who migrated from overseas (Group 1). For Group 3B 23% of its members were aged 20-24 years as of 2001, 23% were aged 20-24 years and 13% were aged 30-34 years. New entrants in the 12 months before the census were generally younger than those LS members entering before this date and much younger than continually resident consistent LS members. In total, 55% of LS members entering in the 12 months before the 2001 census were aged 15-29 years as of April 2001, compared with 25% of LS members who entered between 1991 and 2001 and 24% for LS members who were resident between 1991 and 2001.

Figure 7.2 reflects the empirical regularity in the age pattern at migration. Indeed, research has identified the persistent regularity of such an age-profile internationally (Rogers and Castro, 1981). The highest rates of migration will be among those aged under 30 years and the focus here on those overseas 12 months before will mean that the age pattern among these migrants has probably been identified more clearly. The standard patterning found by Rogers and Castro (1981) arises from the high rates of migration among young adults and young children (who travel with young adults) (Raymer and Rogers, 2008).

Table 7.7: Numbers of ONS LS members per residence pattern by age group as of 2001 census

Group		Age as of 2001 census													Total
		0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60+	
1	LS members at the 2001 census who gave an address a year before as overseas	65	65	38	90	226	164	95	53	32	25	8	7	7	875
2	LS members continually resident, consistent cases 1991-2001	-	-	15,162	14,125	13,038	14,255	16,837	18,636	17,408	16,113	17,922	14,768	11,059	169,323
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	178	404	422	340	792	1,273	993	645	368	213	148	92	49	5,917
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	44	30	24	60	153	151	87	56	21	13	13	5	5	662
Total		287	499	15,646	14,615	14,209	15,843	18,012	19,390	17,829	16,364	18,091	14,872	11,120	132,694

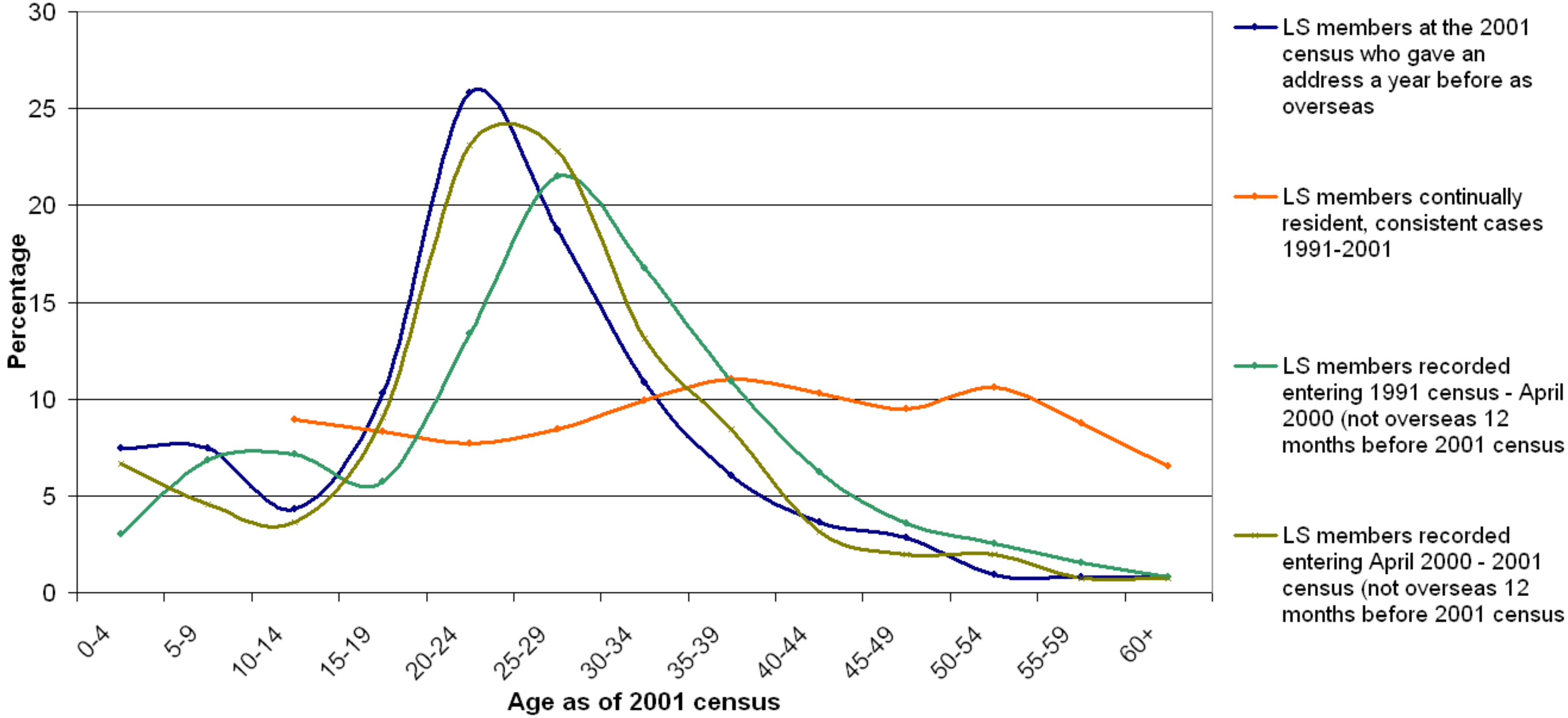
Own elaboration based on ONS LS, June 2011.

Table 7.8: Percentages of ONS LS members per residence pattern by age group as of 2001 census

Group		Age as of 2001 census													Total
		0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60+	
1	LS members at the 2001 census who gave an address a year before as overseas	7.4	7.4	4.3	10.3	25.8	18.7	10.9	6.1	3.7	2.9	0.9	0.8	0.8	100
2	LS members continually resident, consistent cases 1991-2001	-	-	9.0	8.3	7.7	8.4	9.9	11.0	10.3	9.5	10.6	8.7	6.5	100
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	3.0	6.8	7.1	5.7	13.4	21.5	16.8	10.9	6.2	3.6	2.5	1.6	0.8	100
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	6.6	4.5	3.6	9.1	23.1	22.8	13.1	8.5	3.2	2.0	2.0	0.8	0.8	100
Total		0.2	0.4	11.8	11.0	10.7	11.9	13.6	14.6	13.4	12.3	13.6	11.2	8.4	100

Own elaboration based on ONS LS, June 2011.

Figure 7.2: Numbers of ONS LS members per residence pattern by age group as of 2001 census



Own elaboration based on ONS LS, June 2011.

7.4.3 Numbers of LS members by country of birth

Next are the number of new entrants to the LS by the country of birth as recorded at the 2001 census. Tables 7.9 and 7.10 show the numbers and percentages of LS members who were resident in the LS at the 2001 census. Grouped countries used here correspond with those used by the ONS in the FM1 fertility statistics series.

Not surprisingly, the majority of LS members who were continuously resident between 1991 and 2001 are UK born (93%). Among the migrant groups the percentages are much lower; 9% of the sample in Group 1 is from the UK, 12% of Group 3A and 8% of Group 3B. The reason for some of the members of Group 1 coming from the UK is because the LS records migrants from Northern Ireland and Scotland. Among those LS members who were located overseas 12 months before the 2001 census (Group 1) the greatest proportions are from Other European Union (19%) and Australia, New Zealand and Canada (9%). In Group 3A and 3B there are relatively high percentages of LS members from Other European Union. For Group 1 Indian born LS female members make up 5% of the sample, while in Group 3A and 3B they form 4% and 9% respectively. In Group 2 just 1% of the sample was born in India. Among the Pakistani born women most are in Groups 3A (6%) and 3B (6%) with some in Group 1 (3%). For Bangladeshi born women there are more in Groups 3A (4%) and 3B (4%) than Group 1 (1%). This is true also for East African born women – for this category most are in Groups 3A (3%) and 3B (2%) compared with Group 1 (1%).

For all the migrant Groups (1, 3A, 3B) there are high percentages of female LS members who were born in an alternative country. The Other category in Group 1 is 50%, for Group 3A is 48% and for Group 3B is 54%.

Table 7.9: Numbers of ONS LS members per residence pattern by country of birth (grouped) as of 2001 census

	Group	UK	Ireland	Australia, Canada, New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union	Other	Total
1	LS members at the 2001 census who gave an address a year before as overseas	75	22	78	45	30	12	11	170	432	875
2	LS members continually resident, consistent cases 1991-2001	157,146	1,230	379	1,923	1,047	628	843	1,411	4,814	169,421
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	708	160	253	420	349	207	159	804	2,857	5,917
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	55	13	56	25	41	26	16	75	355	662
Total		157,984	1,425	766	2,413	1,467	873	1,029	2,460	8,458	176,875

Own elaboration based on ONS LS, June 2011.

Table 7.10: Percentages of ONS LS members per residence pattern by country of birth (grouped) as of 2001 census

	Group	UK	Ireland	Australia, Canada, New Zealand	India	Pakistan	Bangladesh	East Africa	Other European Union	Other	Total
1	LS members at the 2001 census who gave an address a year before as overseas	8.6	2.5	8.9	5.1	3.4	1.4	1.3	19.4	49.4	100
2	LS members continually resident, consistent cases 1991-2001	92.8	0.7	0.2	1.1	0.6	0.4	0.5	0.8	2.8	100
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	12.0	2.7	4.3	7.1	5.9	3.5	2.7	13.6	48.3	100
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	8.3	2.0	8.5	3.8	6.2	3.9	2.4	11.3	53.6	100
Total		89.3	0.8	0.4	1.4	0.8	0.5	0.6	1.4	4.8	100

Own elaboration based on ONS LS, June 2011.

7.4.4 Numbers of LS members by student status

An important consideration in selecting a sample for fertility analysis is whether the LS member was a student as of the 2001 census. This is important because those LS members in the UK who were students as of the 2001 census would be less likely to give birth and may constitute relatively short-term migrants in England and Wales. The 2001 census variable ('STUP0') is used to identify students. In addition to this, there is the potential for higher levels of attrition for students from overseas (see section 4.5 in Chapter 4).

Tables 7.11 and 7.12 show the numbers and percentages of students as of the 2001 census by type of LS member. As would be anticipated given their older overall age structure, LS members who were continuously resident from 1991 to 2001 have a lower percentage of students as of the 2001 census (17%). Table 7.12 shows that the groups who migrated during the 1990s have higher percentages of school children and students in full-time education as of the 2001 census. Among those LS members who were overseas 12 months before the 2001 census, 40% of LS members were students at the 2001 census. In contrast, and likely to be related to the higher age profile at the 2001 census, 29% of the new entrants between 1991 and 2001 were students as of the 2001 census.

Tables 7.13 and 7.14 present the same analysis of students and non students but only using LS members in the groups who were aged 15-49 years as of the 2001 census. For both the recent migrant Groups 1 and 3B there is a decrease in the percentage who were students when we only use those who were 15-49 years. The largest decrease is among the 3A 1991-2000 migrants group where the percentage of students changes from 29% with no age restriction to 18% when the 15-49 years filter is applied. Among the migrant Group (1) there is a modest decrease in the percentage who are students by about 5%.

Table 7.11: Numbers of ONS LS members per residence pattern by student status at 2001 census

Group		Student status at 2001 census: Schoolchild or student in full-time education	Student status at 2001 census: Not a schoolchild or student in full-time education	Total
1	LS members at the 2001 census who gave an address a year before as overseas	346	529	875
2	LS members continually resident, consistent cases 1991-2001	29,352	140,067	169,419
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	1,725	4,192	5,917
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	198	464	662
Total		31,621	145,252	176,873

Own elaboration based on ONS LS, June 2011.

Table 7.12: Percentages of ONS LS members per residence pattern by student status at 2001 census

Group		Student status at 2001 census: Schoolchild or student in full-time education	Student status at 2001 census: Not a schoolchild or student in full-time education	Total
1	LS members at the 2001 census who gave an address a year before as overseas	39.5	60.5	100
2	LS members continually resident, consistent cases 1991-2001	17.3	82.7	100
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	29.2	70.8	100
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	29.9	70.1	100
Total		17.9	82.1	100

Own elaboration based on ONS LS, June 2011.

Table 7.13: Numbers of ONS LS members per residence pattern by student status at 2001 census – LS members aged 15-49 only

Group		Student status at 2001 census: Schoolchild or student in full-time education	Student status at 2001 census: Not a schoolchild or student in full-time education	Total
1	LS members at the 2001 census who gave an address a year before as overseas	234	451	685
2	LS members continually resident, consistent cases 1991-2001	13,970	96,441	110,411
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001	813	3,811	4,624
3B	census (not overseas 12 months before 2001 census)	129	412	541
Total		15,146	101,115	116,261

Own elaboration based on ONS LS, June 2011.

Table 7.14: Percentages of ONS LS members per residence pattern by student status at 2001 census – LS members aged 15-49 only

Group		Student status at 2001 census: Schoolchild or student in full-time education	Student status at 2001 census: Not a schoolchild or student in full-time education	Total
1	LS members at the 2001 census who gave an address a year before as overseas	34.2	65.8	100
2	LS members continually resident, consistent cases 1991-2001	12.7	87.3	100
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001	17.6	82.4	100
3B	census (not overseas 12 months before 2001 census)	23.8	76.2	100
Total		13.0	87.0	100

Own elaboration based on ONS LS, June 2011.

7.4.5 Numbers of LS members by marital status

Marital status as of the 2001 census for LS members in each of the four groups is shown in Tables 7.15 and 7.16. In keeping with the younger age profile, the Group 1 migrants have a higher percentage of LS members who were single as of the 2001 census compared with all other groups. This includes the LS members entering just before the 2001 census in Group 3B. Interestingly, despite their higher average age, the continually resident consistent cases (Group 2) do not have the highest percentage of LS members married as of the 2001 census. It is the LS migrants entering between the 1991 census and April 2000 who have the highest percentage of LS members who are married.

Table 7.15: Numbers of ONS LS members per residence pattern by marital status at 2001 census

Group		Single (never married)	Married (first marriage)	Re-married	Separated (but still legally married)	Divorced	Widowed	Total
1	LS members at the 2001 census who gave an address a year before as overseas	554	258	27	7	18	11	875
2	LS members continually resident, consistent cases 1991-2001	64,215	68,930	12,183	4,553	15,559	3,979	169,419
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	2,714	2,525	203	210	179	85	5,916
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	349	275	14	6	13	5	662
Total		67,832	71,988	12,427	4,776	15,769	4,080	176,872

Own elaboration based on ONS LS, June 2011.

Table 7.16: Percentages of ONS LS members per residence pattern by marital status at 2001 census

Group		Single (never married)	Married (first marriage)	Re-married	Separated (but still legally married)	Divorced	Widowed	Total
1	LS members at the 2001 census who gave an address a year before as overseas	63.3	29.5	3.1	0.8	2.1	1.3	100
2	LS members continually resident, consistent cases 1991-2001	37.9	40.7	7.2	2.7	9.2	2.3	100
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	45.9	42.7	3.4	3.5	3.0	1.4	100
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	52.7	41.5	2.1	0.9	2.0	0.8	100
Total		38.4	40.7	7.0	2.7	8.9	2.3	100

Own elaboration based on ONS LS, June 2011.

7.4.6 Numbers of LS members by Government Office region

Tables 7.17 and 7.18 provide an overview of the geography of the recent migrant and non-migrant comparison groups. London, the South East and the East of England are the areas where most of the LS members of the migrant group were resident as of the 2001 census. The North West and the West Midlands have moderate shares and all other areas fewer LS members of this type. London has the largest share – 28% of LS members in Group 1 were living in the capital. Among the continually resident consistent cases there is a more even geographic spread of female LS members. Groups 3A and 3B show many similarities in their percentages, London has an even greater share in these groups with 45% of LS members in both categories living in London as of the 2001 census. Interestingly, when compared to Group 1, the only other area with a relatively high percentage of LS members is the South East.

Table 7.17: Numbers of ONS LS members per residence pattern by Government Office region at 2001 census

Group		North East	Wales	North West	Yorkshire and The Humber	East Midlands	West Midlands	East of England	London	South East	South West	Total
1	LS members at the 2001 census who gave an address a year before as overseas	26	22	76	79	43	47	107	242	181	52	875
2	LS members continually resident, consistent cases 1991-2001	8,577	9,494	22,440	17,025	14,382	17,921	17,718	19,249	26,211	16,347	169,364
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	107	119	405	297	282	421	439	2,635	897	305	5,907
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	18	10	36	48	28	46	52	298	99	26	661
Total		8,728	9,645	22,957	17,449	14,735	18,435	18,316	22,424	27,388	16,730	176,807

Own elaboration based on ONS LS, June 2011.

Table 7.18: Percentages of ONS LS members per residence pattern by Government Office region at 2001 census

Group		North East	Wales	North West	Yorkshire and The Humber	East Midlands	West Midlands	East of England	London	South East	South West	Total
1	LS members at the 2001 census who gave an address a year before as overseas	3.0	2.5	8.7	9.0	4.9	5.4	12.2	27.7	20.7	5.9	100
2	LS members continually resident, consistent cases 1991-2001	5.1	5.6	13.2	10.1	8.5	10.6	10.5	11.4	15.5	9.7	100
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	1.8	2.0	6.9	5.0	4.8	7.1	7.4	44.6	15.2	5.2	100
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	2.7	1.5	5.4	7.3	4.2	7.0	7.9	45.1	15.0	3.9	100
Total		4.9	5.5	13.0	9.9	8.3	10.4	10.4	12.7	15.5	9.5	100

Own elaboration based on ONS LS, June 2011.

7.4.7 Numbers of LS members by occupation status

Tables 7.19 and 7.20 show the numbers and percentages of LS members by occupation status at the 2001 census. Overall, most LS members, regardless of the group they are in, fall into the Professional Occupations group (39%). However, there are big variations between the different groups – the largest percentage of the Professional Occupations are those persons in Group 3A (3.1%). The Skilled non-manual occupations category has the second largest overall percentage of LS members (31%). Remarkable is the way in which there are so few LS members from Group 2 which are mainly concentrated in the Not classified type (97%). In the Not classified group there are smaller percentages from the other groups – just 2% from Group 1 and 3A and 7% from Group 3B. With so many of the LS members located in this group the usefulness of this covariate is questionable and perhaps another variable should be used in subsequent analysis.

Table 7.19: Numbers of ONS LS members per residence pattern by occupation status at 2001 census (ages 15-49 at 2001 census)

	Group	Not classified - occupation code or employment status not available	I Professional occupations	II Managerial and technical occupations	IIIN Skilled non-manual occupations	IIIM Skilled manual occupations	IV Partly-skilled occupations	V Unskilled occupations	Occupation inadequately described	Total
1	LS members at the 2001 census who gave an address a year before as overseas	192	435	2,934	2,964	819	1,642	389	171	9,546
2	LS members continually resident, consistent cases 1991-2001	14,098	41	172	132	18	97	21	12	14,591
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	1,619	2,991	28,333	34,876	7,808	16,989	4,490	725	97,831
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	241	291	1,028	830	182	472	120	72	3,236
Total		16,150	3,758	32,467	38,802	8,827	19,200	5,020	980	125,204

Own elaboration based on ONS LS, August 2011.

Table 7.20: Percentages of ONS LS members per residence pattern by occupation status at 2001 census (ages 15-49 at 2001 census)

	Group	Not classified - occupation code or employment status not available	I Professional occupations	II Managerial and technical occupations	IIIN Skilled non-manual occupations	IIIM Skilled manual occupations	IV Partly-skilled occupations	V Unskilled occupations	Occupation inadequately described	Total
1	LS members at the 2001 census who gave an address a year before as overseas	2.0	4.6	30.7	31.0	8.6	17.2	4.1	1.8	100
2	LS members continually resident, consistent cases 1991-2001	96.6	0.4	1.2	0.9	0.1	0.7	0.1	0.1	100
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	1.7	31.3	29.0	35.6	8.0	17.4	4.6	0.7	100
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	7.4	3.0	31.8	25.6	5.6	14.6	3.7	2.2	100
Total		12.9	39.4	25.9	31.0	7.1	15.3	4.0	0.8	100

Own elaboration based on ONS LS, August 2011.

7.4.8 Conclusions

Covariate information from the 2001 census provides detail on the characteristics of the samples identified and is useful with a view to devising regression a model for duration to first birth allowing estimation of the relative risk of birth. The number of LS members entering in the 12 months before the census (including those from overseas) is very slightly higher than the average per annum figures for the 1991 – April 2000 period.

Analysis of the ages of LS members in the entry forms identified showed that new entrants in the 12 months before the 2001 census were generally younger than entrants between 1991 and 2001. The LS members in Group 1 are predominantly in the 20-24 years age group; this may at least in part be related to the narrower time-frame which has been selected and the proximity to the 2001 census, which was where the age variable in use was selected. This would make the sample seem younger, relative to new entrants between 1991 and 2001. Consistent with this younger age profile, the Group 1 migrants have a higher percentage of LS members who were single as of the 2001 census compared with all other groups. Despite the higher average age of the continually resident consistent cases, it is the LS migrants entering between the 1991 census and April 2000 who have the highest percentage of LS members who are married.

Among LS members at the 2001 census and giving an overseas address 12 months previously, most came from within the European Union (as of 2001) and East Asia. However, most LS members in Group 1 moved from an 'other' country. Compared to the other groups, for the migrants at the 2001 census a higher percentage were students or school children (the percentage dropped slightly when the 15-49 age range was used). Occupational status of the LS members as of the 2001 census was relatively uninformative with high proportions of the LS members falling into the 'Not classified' category.

7.5 Embarkations and deaths (2001-2004) among migrants and comparison groups at the 2001 census

As already identified in this thesis, Chapters 3 and 4 discussed the way in which the ONS LS records embarkations from England and Wales through using NHSCR records on GP de-registrations and deaths. In order to identify the sample of LS migrants which can be used in event history analysis it is necessary to identify where there were embarkations and deaths among LS members after the 2001 census. In this section the recorded embarkation and death information will be considered. By using this information important detail on the groups identified for analysis is available and the attrition of LS members after the 2001 census is clear.

7.5.1 Post-2001 census embarkations

For LS members at the 2001 census the date of first embarkation from the NHSCR is used to identify where there was a departure. (Analysis for subsequent embarkations is not possible because of small numbers which cannot be disclosed). The reason for completing this analysis is to give an indication of the number of LS members who leave the LS and hence need to be censored in the period of interest.

Tables 7.21, 7.22 and 7.23 show the numbers and percentages of LS members at the 2001 census and subsequent embarkations. For the new entrants from overseas at the 2001 census (Group 1) there are a total of 59 LS members who are recorded as having left the LS in the 2001-2008 period. In total, this represents 7% of the sample of LS members of this type at the 2001 census (875). There are slightly more LS members of this type lost in the early part of the time-frame (2001-2004) compared with the latter part where, in the years 2005-2008, there are less than 10 members embarking each year. Some of this could be related to the student status as of the 2001 census (students returning to their country of origin after finishing studies). The continually resident consistent cases in the 1991-2001 period have quite the opposite pattern, with increasing numbers embarking with time from 2001. The group of LS members entering the LS between April 2000 and the 2001 census show a slight increase in embarkations around the years 2002-2004, after which there is again a decrease.

Table 7.21: Number of ONS LS members per residence pattern by first embarkation (2001-2008) (ages 15-49 at 2001 census)

Group		First embarkation								Total
		2001	2002	2003	2004	2005	2006	2007	2008	
1	LS members at the 2001 census who gave an address a year before as overseas	11	9	10	13	*	5	6	*	54
2	LS members continually resident, consistent cases 1991-2001	33	63	44	85	76	103	102	93	599
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	16	33	35	26	27	22	22	13	194
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	4	*	5	4	*	4	*	*	17
Total		64	105	94	128	103	134	130	106	864

*Own elaboration based on ONS LS, August 2011. Note: * denotes values have been deleted to meet ONS disclosure controls.*

Table 7.22: Percentage of ONS LS members per residence pattern by first embarkation (2001-2008) (ages 15-49 at 2001 census) (of all embarkations and of all LS members in each group)

Group		First embarkation								% group	Total
		2001	2002	2003	2004	2005	2006	2007	2008		
1	LS members at the 2001 census who gave an address a year before as overseas	20.4	16.7	18.5	24.1	*	9.3	11.1	*	6.2	100
2	LS members continually resident, consistent cases 1991-2001	5.5	10.5	7.3	14.2	12.7	17.2	17.0	15.5	0.4	100
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	8.2	17.0	18.0	13.4	13.9	11.3	11.3	6.7	3.3	100
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	23.5	*	29.4	23.5	*	23.5	*	*	2.6	100
Total		7.4	12.2	10.9	14.8	11.9	15.5	15.0	12.3	0.5	100

*Own elaboration based on ONS LS, August 2011. Note: * denotes values have been deleted to meet ONS disclosure controls.*

Table 7.23: Number of ONS LS members per residence pattern by first embarkation (2001-2008) (ages 15-49 and not students as of the 2001 census)

Group		First embarkation								Total
		2001	2002	2003	2004	2005	2006	2007	2008	
1	LS members at the 2001 census who gave an address a year before as overseas	4	7	8	9	*	4	3	*	35
2	LS members continually resident, consistent cases 1991-2001	30	47	42	78	66	93	87	82	525
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	11	25	26	21	23	19	21	11	157
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	*	*	5	*	*	3	*	*	8
Total		45	79	81	108	89	119	111	93	725

*Own elaboration based on ONS LS, August 2011. Note: * denotes values have been deleted to meet ONS disclosure controls.*

Table 7.24: Percentage of ONS LS members per residence pattern by first embarkation (2001-2008) (ages 15-49 and not students as of the 2001 census) (of all embarkations and of all LS members in each group)

Group		First embarkation								% group	Total
		2001	2002	2003	2004	2005	2006	2007	2008		
1	LS members at the 2001 census who gave an address a year before as overseas	11.4	20.0	22.9	25.7	*	11.4	8.6	*	4.0	100
2	LS members continually resident, consistent cases 1991-2001	5.7	9.0	8.0	14.9	12.6	17.7	16.6	15.6	0.3	100
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	7.0	15.9	16.6	13.4	14.6	12.1	13.4	7.0	2.7	100
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	*	*	62.5	*	*	37.5	*	*	1.2	100
Total		6.2	10.9	11.2	14.9	12.3	16.4	15.3	12.8	0.4	100

*Own elaboration based on ONS LS, August 2011. Note: * denotes values have been deleted to meet ONS disclosure controls.*

7.5.2 Deaths after the 2001 census

Using the date of death for LS members, the number of deaths per year after the 2001 census can be calculated for the trajectories identified at the 2001 census. Tables 7.25 and 7.26 show the numbers and percentages of deaths between 2001 and 2007. This is important as it gives an indication of the number of LS members for each type at the 2001 census which are lost from the 2001-2007 period and therefore may not be used in subsequent analysis. For the migrant group (1) there are no values because there are a small number of deaths which means the values cannot be disclosed. The higher overall age profile among LS members who were continuously resident consistent cases between 1991 and 2001 and the large number of LS members in this group is probably leading to the higher number of deaths observed. There are relatively small numbers of deaths in the 1991-2001 new entrants group, again leading to suppression of the values.

Table 7.25: Number of ONS LS members per residence pattern who die (2001-2007) (ages 15-49 at 2001 census)

Group		Year of death							Total
		2001	2002	2003	2004	2005	2006	2007	
1	LS members at the 2001 census who gave an address a year before as overseas	*	*	*	*	*	*	*	*
2	LS members continually resident, consistent cases 1991-2001	17	9	11	9	15	17	18	96
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	43	86	103	99	92	100	108	631
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	*	6	*	3	4	4	*	17
Total		60	101	114	111	111	121	126	744

*Own elaboration based on ONS LS, August 2011. Note: * denotes values have been deleted to meet ONS disclosure controls.*

Table 7.26: Percentage of ONS LS member per residence pattern who die (2001-2007) (of all deaths) (ages 15-49 at 2001 census)

Group		Year of death							Total
		2001	2002	2003	2004	2005	2006	2007	
1	LS members at the 2001 census who gave an address a year before as overseas	*	*	*	*	*	*	*	*
2	LS members continually resident, consistent cases 1991-2001	17.7	9.4	11.5	9.4	15.6	17.7	18.8	100
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	6.8	13.6	16.3	15.7	14.6	15.8	17.1	100
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	*	35.3	*	17.6	23.5	23.5	*	100
Total		8.1	13.6	15.3	14.9	14.9	16.3	16.9	100

*Own elaboration based on ONS LS, August 2011. Note: * denotes values have been deleted to meet ONS disclosure controls.*

Table 7.27: Percentage of ONS LS members per residence pattern who die (2001-2007) (of all LS members in each group) (ages 15-49 at 2001 census)

Group		Year of death							Total
		2001	2002	2003	2004	2005	2006	2007	
1	LS members at the 2001 census who gave an address a year before as overseas	*	*	*	*	*	*	*	*
2	LS members continually resident, consistent cases 1991-2001	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1
3A	LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)	0.7	1.5	1.7	1.7	1.6	1.7	1.8	10.7
3B	LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)	*	0.9	*	0.5	0.6	0.6	*	2.6
Total		0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.4

*Own elaboration based on ONS LS, August 2011. Note: * denotes values have been deleted to meet ONS disclosure controls.*

7.5.3 Conclusions on right censoring – post-2001 departures

Identifying what happens to the groups identified at the 2001 census is important for sample selection for event history analysis, data on embarkations and deaths is discussed in this subsection. Among the migrant group (Group 1) there are a total of 59 LS members who leave in the 2001-2007 period. Most of these departures are in the earlier part of the period (2002-2004). As suggested earlier, this could be related to the student status as of the 2001 census (students returning to their country of origin after finishing studies) and this should be explored further and considered by age group. Subsequent embarkations cannot be included because the values are too small to be cleared by the ONS. Deaths among the samples selected are also relatively small and concentrated in the continually consistent cases (which has an older age profile) and Group 3A, which is the largest migrant comparison group.

7.6. Fertility (1999-2004) among migrants and comparison groups at the 2001 census

With the groups as of the 2001 census identified it is possible to calculate age-specific fertility rates (ASFRs) and total fertility rates (TFRs) for the years after the 2001 census. This section presents fertility rates for the years immediately after the 2001 census. The purpose of this is to give an indication of the overall level of fertility among the different groups and to identify if there seems to be a link between migration and elevated fertility.

7.6.1 Fertility rates for 2001-2004

Tables 7.28 to 7.31 contain small numbers of women and births which lead to fertility rates which in some cases can be distorted by the small counts. Group 3B in the series, 'LS members recorded entering April 2000 – 2001 census' have the highest fertility in 2001 and throughout the series of years. In 2001 the TFR is 4.77 which is remarkably high and the result of very high ASFRs, especially in the 30-34 age group. It is possible that the elevated fertility which is shown in this series of values is related to the registration with a GP in the 12 months before the 2001 census. According to the census response these LS members were not overseas 12 months before the 2001 census. After the much higher TFR in 2001 there is a decline to a TFR around 2 in 2002 and then 2.6 and 2.2 in 2003 and 2004 respectively.

In this chapter Group 1 has been identified as the set of recent migrants and there is a decrease in the TFR from 2001 – the rate is 2.1 in 2001 before dropping to a rate of 1.1 in the other years in the period. The number of women resident in each of the years remains roughly the same. However, the number of births decreases slightly from 2001, leading to the lower TFRs. The analysis is failing to capture all woman who are leaving (some will leave via attrition) which means that the denominator (the number of women) is inflated relative to the numerator. Among the continually resident consistent cases (Group 2) there are larger counts of women and births. The rates which are derived are fairly consistent year-to-year.

Group 3A are those LS members who migrated to England and Wales between 1991 and 2000. This has a high level of fertility compared to the continually resident consistent cases and the recent migrants as of the 2001 census. Interestingly, among this specific group the rate is consistently high, although it does decrease a little from 2001 and 2002 to 2003 and 2004.

Group 3B is a group which was selected based on the registration with a GP in the April 2000-April 2001 period and being resident at the 2001 census. Importantly, these LS members reported on the 2001 census form that they were not living overseas 12 months before the 2001 census (April 2000). This indicates that there was a lag between entry to England and Wales (sometime before April 2000) and registration with a GP (sometime between April 2000 and April 2001 (the census)). This suggests that the TFR for this group in 2001 (4.7) could be a result of the association between the timing of GP registration (as previously discussed in Chapter 6) and becoming pregnant. The higher rates (than other groups) in the years 2003 and 2004 might be related to the migration event and constitute the 'real' post migration fertility trend.

Table 7.28: Women, births and ASFRs for 2001 by group at the 2001 census

Age	1- LS members at the 2001 census who gave an address a year before as overseas			2- LS members continually resident, consistent cases 1991-2001			3A- LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)			3B- LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)		
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000
15-19	51	4	78.4	14,474	350	24.2	325	10	30.8	49	7	142.9
20-24	230	15	65.2	13,404	794	59.2	706	77	109.1	142	20	140.8
25-29	171	10	58.5	13,616	1,292	94.9	1,234	139	112.6	164	30	182.9
30-34	111	8	72.1	16,666	1,420	85.2	1,056	123	116.5	93	22	236.6
35-39	49	6	122.4	18,365	680	37.0	700	36	51.4	61	10	163.9
40-44	*	*	<30.0	17,567	121	6.9	387	6	15.5	*	*	<100.0
45-49	*	*	<0.5	16,135	7	0.4	230	*	4.3	*	*	<0.5
TFR			2.12			1.54			2.20			4.77

*Own elaboration based on ONS LS, June 2011. Note: * denotes values have been deleted to meet ONS disclosure controls.*

Table 7.29: Women, births and ASFRs for 2002 by group at the 2001 census

Age	1- LS members at the 2001 census who gave an address a year before as overseas			2- LS members continually resident, consistent cases 1991-2001			3A- LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)			3B- LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)		
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000
15-19	*	*	<0.5	14,703	292	19.9	351	8	22.8	32	3	93.8
20-24	196	10	51.0	13,599	770	56.6	593	66	111.3	131	12	91.6
25-29	189	8	42.3	12,890	1,171	90.8	1,120	112	100.0	165	10	60.6
30-34	127	9	70.9	15,990	1,350	84.4	1,139	120	105.4	111	8	72.1
35-39	*	*	<30.0	18,214	685	37.6	747	62	83.0	61	3	49.2
40-44	*	*	<30.0	17,915	141	7.9	441	8	18.1	*	*	<50.0
45-49	*	*	<0.5	16,357	7	0.4	*	*	<0.5	*	*	<0.5
TFR			1.05			1.49			2.20			2.01

*Own elaboration based on ONS LS, June 2011. Note: * denotes values have been deleted to meet ONS disclosure controls.*

Table 7.30: Women, births and ASFRs for 2003 by group at the 2001 census

Age	1- LS members at the 2001 census who gave an address a year before as overseas			2- LS members continually resident, consistent cases 1991-2001			3A- LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)			3B- LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)		
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000
15-19	*	*	<0.5	15,041	377	25.1	355	10	28.2	*	*	<0.5
20-24	149	10	67.1	13,501	797	59.0	495	48	97.0	113	14	123.9
25-29	208	12	57.7	12,609	1,114	88.3	1,019	105	103.0	168	17	101.2
30-34	141	6	42.6	15,501	1,348	87.0	1,203	117	97.3	118	15	127.1
35-39	65	3	46.2	17,881	811	45.4	809	48	59.3	63	11	174.6
40-44	*	*	<0.5	18,102	142	7.8	496	7	14.1	*	*	<0.5
45-49	*	*	<0.5	16,719	6	0.4	*	*	<0.5	*	*	<0.5
TFR			1.07			1.56			1.99			2.63

*Own elaboration based on ONS LS, June 2011. Note: * denotes values have been deleted to meet ONS disclosure controls.*

Table 7.31: Women, births and ASFRs for 2004 by group at the 2001 census

Age	1- LS members at the 2001 census who gave an address a year before as overseas			2- LS members continually resident, consistent cases 1991-2001			3A- LS members recorded entering 1991 census - April 2000 (not overseas 12 months before 2001 census)			3B- LS members recorded entering April 2000 - 2001 census (not overseas 12 months before 2001 census)		
	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000	Women	Births	Rate per 1,000
15-19	*	*	<0.5	15,058	374	24.8	374	3	8.0	*	*	<50.0
20-24	123	4	32.5	13,717	848	61.8	395	31	78.5	82	14	170.7
25-29	199	17	85.4	12,622	1,064	84.3	908	104	114.5	168	11	65.5
30-34	144	8	55.6	14,963	1,494	99.8	1,241	119	95.9	130	11	84.6
35-39	75	3	40.0	17,629	780	44.2	890	43	48.3	71	5	70.4
40-44	*	*	<0.5	18,363	164	8.9	535	14	26.2	*	*	<0.5
45-49	*	*	<0.5	16,967	6	0.4	*	*	<10.0	*	*	<0.5
TFR			1.07			1.62			1.89			2.16

*Own elaboration based on ONS LS, June 2011. Note: * denotes values have been deleted to meet ONS disclosure controls.*

7.6.2 Conclusions on fertility rates around the 2001 census

To try and identify if there is a migration and fertility linkage this section has calculated fertility rates for the 2001-2004 period. In many cases the tables here have included small numbers and hence interpretation must be cautious. For the 2001-2004 period there are some interesting findings in relation to the fertility for the groups selected. Among Group 1 there is a higher level of fertility in 2001 (2.1) relative to the subsequent years (1.1 for years 2002-2004). This suggests that there may be a link between movement to England and Wales and subsequently higher fertility. This finding is of particular interest as this is independent of the timing of arrival in England and Wales – the date of registration with a GP has not been used and instead the self-reported information on where the LS member was resident 12 months before the census has been used. As a comparison, Group 3B has the highest fertility throughout the period 2001-2004 with a TFR in 2001 of 4.77 which declines sharply thereafter. As discussed in the text above, this group includes the date of registration on the NHSCR so it is possible that this is a result of the registration effect identified previously (see Chapter 6). Given the insights related to the attrition of foreign born women which were made in Chapter 4 (they were found to be three times more likely to drop out between census dates) it is also possible that the denominator is over-inflated with increasing duration from the census. This may lead to a population exposed to risk which is too large relative to the numbers of women who are actually resident and give birth.

7.6.3 Fertility analysis for migrants arriving within 12 months of the 1991 census

In order to start to account for attrition among migrants at the 2001 census it is possible to replicate the analysis for the years after 2001 but use those migrants who arrived in the 12 months before 1991 and were also resident at the 2001 census. Interpretation of the fertility for migrants who arrived before 2001 can then be considered alongside fertility measures which consider attrition among migrants from the 1991 census. The rationale is that women arriving in the year before 1991 and at the 2001 census are a sample who in 1991, 1992 and 1993 were recent immigrants and remained resident. It is possible to be reasonably confident that the denominators for this group are correct. The same cannot be said for the sample arriving before the 2001 census. If the pattern among the group of migrants for the years 1991, 1992 and 1993 is similar to that for 2001, 2002 and 2003 it is possible to be confident that emigration is not biasing results.

LS members who migrated within 12 months of the 1991 census and still resident as of the 2001 census were selected using the comparable indicator on migration status from the 1991 census (MIGPOP9). This group consists of about 450 LS members. Insufficient numbers of women and births were available for the calculation of reliable age-specific fertility rates (or the clearance of these outputs in five year groups from the 'safe setting' at the Office for National Statistics). Instead the general fertility rate (GFR) was calculated.

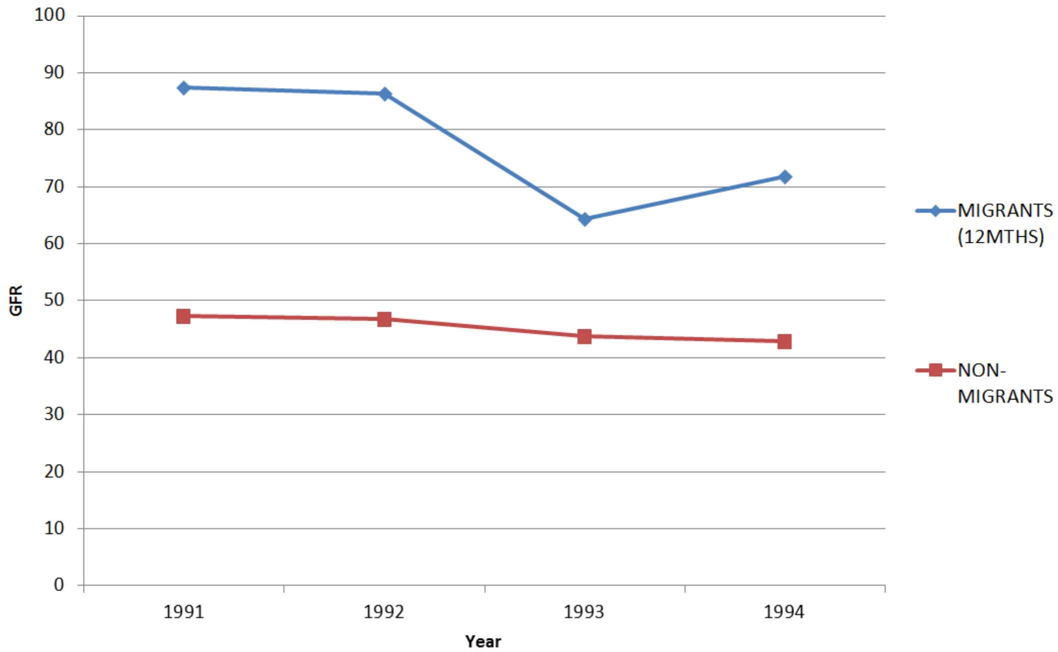
Table 7.32 and Figure 7.3 show the GFR for the migrant and non-migrant groups for 1991-1994. There is a negligible decline in the GFR from 1991 for the non-migrants and a sharper decline for the migrants.

Table 7.32: General fertility rates – 1991-1994

Migrant group	1991	1992	1993	1994
Migrants (12 months)	87.3	86.3	64.3	71.7
Non-migrants	47.4	46.7	43.8	42.9

Own elaboration based on ONS LS, September 2012.

Figure 7.3: General fertility rates – 1991-1994



Own elaboration based on ONS LS, September 2012.

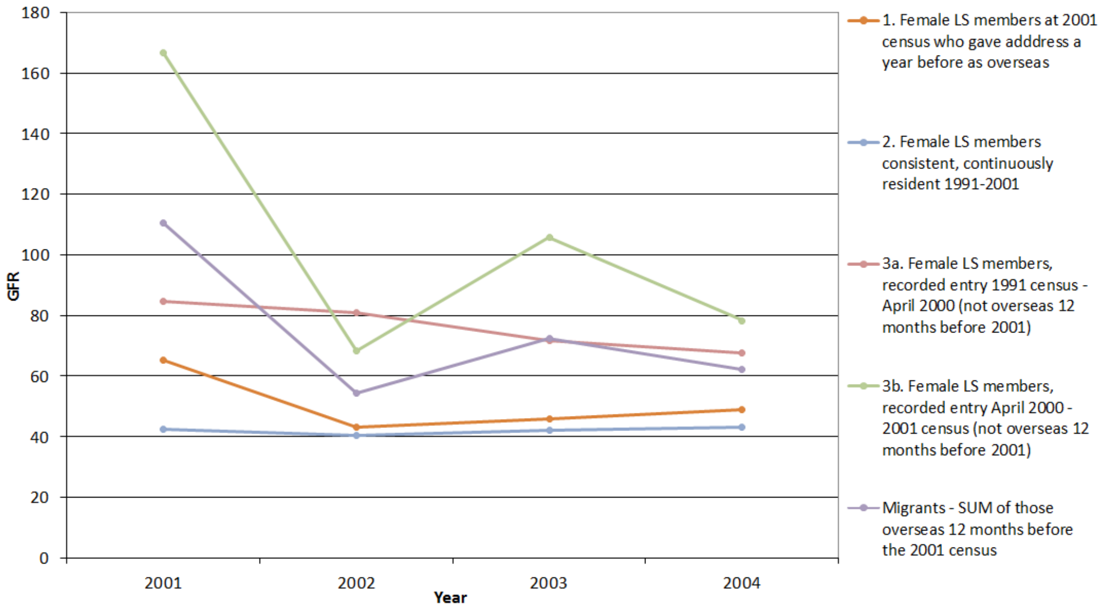
Table 7.33 and Figure 7.4 show corresponding values for 2001-2004 and for the combined 12 month migrant group (see the purple line in the figure). Note that this sample is not selected on the basis of being at the subsequent census (2011, for which there is no data at the present time). Non-migrants (continually resident consistent cases) have a similar level of fertility to the corresponding non-migrant group at the 1991 census.

Table 7.33: General fertility rates – 2001-2004

Migrant group	2001	2002	2003	2004
1. Female LS members at 2001 census who gave adddress a year before as overseas	65.2	43.1	45.9	48.9
2. Female LS members consistent, continuously resident 1991-2001	42.3	40.3	42.0	43.3
3a. Female LS members, recorded entry 1991 census - April 2000 (not overseas 12 months before 2001)	84.5	80.9	71.9	67.7
3b. Female LS members, recorded entry April 2000 - 2001 census (not overseas 12 months before 2001)	166.7	68.4	105.8	78.4
Migrants - SUM of those overseas 12 months before the 2001 census	110.6	54.4	72.5	62.1

Own elaboration based on ONS LS, September 2012.

Figure 7.4: General fertility rates – 2001-2004



Own elaboration based on ONS LS, September 2012.

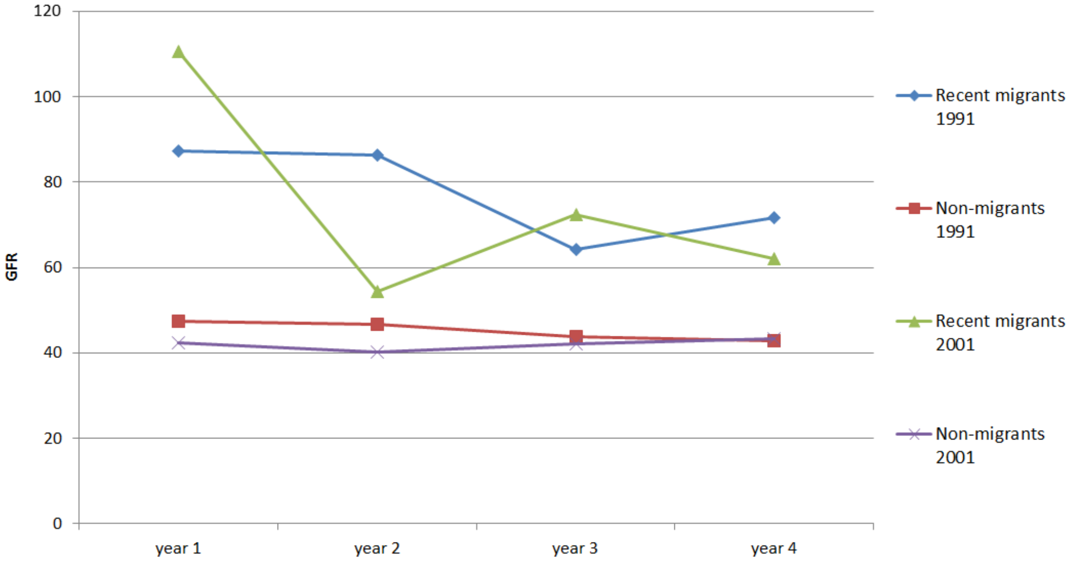
Table 7.34 and Figure 7.5 compare the 1991 and 2001 values. The 2001 migrant group has a higher level of fertility in year 1 compared with the 1991 migrant group but this declines for 1992 before becoming broadly comparable with the 1991 figure for years three and four. It is notable that the 1991-1994 sample has a similar fertility profile despite this being restricted to LS members who were at the 2001 census.

Table 7.34: General fertility rates – 1991-1994 and 2001-2004

Migrant group	year 1	year 2	year 3	year 4
Recent migrants 1991	87.3	86.3	64.3	71.7
Non-migrants 1991	47.4	46.7	43.8	42.9
Recent migrants 2001	110.6	54.4	72.5	62.1
Non-migrants 2001	42.3	40.3	42.0	43.3

Own elaboration based on ONS LS, September 2012.

Figure 7.5: General fertility rates – 1991-1994 and 2001-2004



Own elaboration based on ONS LS, September 2012.

Overall, the results indicate that emigration after 2001 of migrants in the sample who were overseas 12 months before the 2001 census might be affecting the fertility rates derived. This may be the case because the decline in fertility after 2001 is greater than for the corresponding group for 1991-1994 who arrived 12 months before 1991 and were resident at the 2001 census. The decline in fertility for the recent migrants as of the 1991 census suggests that emigration may be impacting on the post-2001 fertility rates which were arrived at (see Figure 7.5 which shows the difference in profiles between the green and blue lines). At the current time, until 2011 census data is available to enable identification of embarkation and attrition it is not possible to identify attrition among the 2001 migrant group. From the analysis it is clear that the fertility of those who arrived in a year before the census is greater than non-migrants.

7.7 Conclusions and implications

This chapter has developed the findings made in previous chapters to select LS members for analysis, based on their form of residence and entry to the LS around the 2001 census. The first half of the analysis in this chapter was focused on identifying a group of LS members who migrated to England and Wales and were resident at the 2001 census. In section 7.3 the key points to consider in the selection of LS members based on their NHSCR registration were outlined, leading to the identification of a group of recent migrants (Group 1) and two comparison groups (continually resident consistent cases (between 1991 and 2001), Group 2) and LS members who entered the LS through registration with a GP sometime between the 1991 and 2001 census (Groups 3A and 3B).

Section 7.4 outlined the socio-economic information for these groups. As would perhaps be anticipated, the recent migrant group at the 2001 census had a younger overall age profile as of the 2001 census compared to the other groups. Fitting with this, the group had a lower percentage of members who were married and a higher percentage who were students. Section 7.5 gave the numbers of post 2001 census embarkations and deaths. Among the migrant group there were 59 embarkations (2001-2007) and this analysis is primarily interested in the 2001-2004 period so it is important to note that most of these embarkations were earlier in the decade (2001-2004 have the highest levels of embarkations). The number of deaths in the recent migrant group was too small to consider. In addition to the recorded embarkations and deaths there will be unrecorded embarkations from the ONS LS which cannot be identified until the 2011 census data is linked and loss to follow-up identified.

In section 7.6 the fertility of the different groups was estimated. For the recent migrants, Group 1, there was a higher level of fertility in 2001 compared to the subsequent years (2002-2004). This is an important finding because the focus of the analysis has been the period immediately after the 2001 census, when attrition is less likely to be making such a large impact on findings and also that the sample which has been selected is one where the duration from migration, although imprecise, does tell us that this was a recent migrant group. In parallel to this, Group 3B, which selects only those migrants who registered with a GP between April 2000 and the 2001 census and were not overseas 12 months before the 2001 census, showed a much higher rate of fertility in 2001, after which there was a decline similar to that in Group 1. However, with this sample the ability to disentangle NHSCR registration because of wishing to give birth and migration and subsequent elevated fertility is more complex. The comparison to a group of recent migrants at the 1991 census has enabled the identification of whether attrition may be an impacting on the results if there are similar levels of emigration or drop out among migrants at the 2001 census. Through using a sample of recent migrants at the 1991 census who were still resident at the 2001 census it has become apparent that there was an elevated level of fertility among recent migrants at the 1991 census and that the rate was higher in year 2 (1992) than for 2002, suggesting that when emigration is taken into account the denominator is reduced.

Despite the approach here of using a sample of recent migrants at the 2001 census, the recent migrant sample identified is not suited for event-history analysis on its own because of its relatively small size. The identification of the different types of migrant in this Chapter and the way in which this selection of the different groups has been developed (with the use of information on the trace and the migration status at the 2001 census) means that migrant groups can be used in subsequent analysis. By taking a different approach to estimating the date of arrival for migrants to England and Wales it will be possible to use the groups identified in subsequent analysis as a covariate. This chapter has added to the present analysis in this thesis by identifying that the sample of migrants at the 2001 census, who arrived at some point in the 1991-2001 period, had a higher rate of fertility around the time of the 2001 census.

The next chapter selects migrants who entered the ONS LS in the 1991-2001 period and were at the 2001 census. Based on the insights made here on the departures of LS members who were migrants between 1991 and 2001, and also the attrition previously discussed, the first two years after the 2001 census will be used in this analysis.

Chapter 8

Estimating the fertility of recent migrants at the 2001 census

Chapter abstract

The previous two chapters have shown that there is systematic bias in the entry of Office for National Statistics (ONS) Longitudinal Study (LS) members and their fertility (Chapter 6) and that there is a relatively small sample of LS members who were resident overseas 12 months before the 2001 census with which to conduct analysis (Chapter 7). Therefore, this chapter uses all female LS members entering the LS (and England and Wales) between the 1991 census and the 2001 census to estimate the risk of birth in the 2001-2003 period. The main question is – what is the fertility of recent migrants to England and Wales relative to non-migrants in the 2001-2003 period? The findings in Chapter 4 on attrition in the 1991-2001 period necessitates working in the two years immediately after the 2001 census as this minimises potential bias to the estimates from unidentifiable attrition.

8.1 Introduction

The previous chapter (Chapter 7) identified that in addition to a relationship between entry to the Office for National Statistics (ONS) Longitudinal Study (LS) and subsequent fertility (see Chapter 6) there is a relatively small sample of migrants who arrived in England and Wales just before the 2001 census and said that they were living overseas 12 months before. In this chapter a sample of LS members who migrated to England and Wales in the period between the 1991 and 2001 census is selected and the fertility of these LS members after the 2001 census estimated.

In order to use the migrants entering the ONS LS in the 1991-2001 period, assumptions about the date of migration to England and Wales are made. This is because, as previously shown, the date of registration on the NHSCR (as recorded in the ONS LS) is not necessarily the same as the actual date of migration; it is possible the migration event was at an earlier time and therefore, this must be factored into estimates of the duration from migration to birth. In the approach used here the date of NHSCR registration will be one of four migration dates used. Subsequently, these four dates are used to calculate the duration to each month in the 24 months after the 2001 census. The four duration assumptions are used in repeated versions of the same discrete time hazards model to estimate the risk of birth. In addition to this new approach of creating continuous measures of duration since migration to England and Wales the migrant groups which were identified and profiled in the last chapter are used to estimate the relative fertility for each group controlling for their socio-demographic characteristics.

In section 8.2 the rationale for this work is detailed before section 8.3 outlines the research questions with which this analysis is concerned. Section 8.4 details the methodology for answering research question one and the samples of migrants and non-migrants to be used in this analysis. In section 8.5 methodology for discrete-time hazards regression models are introduced. Section 8.6 presents descriptive statistics for the migrant and non-migrant groups. A discussion of results for life table analysis for risk of birth for migrants and non-migrants is made in section 8.7.

In section 8.8 results from discrete-time hazards models are shown for several measures of migration before conclusions are drawn in section 8.9.

8.2 Rationale

One of the primary aims of this thesis has been to estimate fertility among migrants to England and Wales relative to their date of migration to England and Wales. Given the problem with the date of NHSCR registration, which the ONS LS suggests is a proxy for migration to England and Wales, an alternative approach is required. This section explains the rationale and approach to be adopted.

The rationale for this work remains as was previously outlined, to estimate the risk of birth among a group of recent migrants to England and Wales and to compare this to non-migrants in the ONS LS. To answer the overarching question about the fertility of migrants to England and Wales it is necessary to consider this as three inter-related research questions. The first of these is simply, is there any difference in fertility between the migrant and non-migrant groups. Based on this, it is necessary to introduce the socio-economic / demographic characteristics of the migrant and non-migrant groups and assess the degree to which these characteristics may be accounting for any difference between the two groups. For example, in Chapter 7 it was identified that the migrant group matched the typical age at migration schedule (Rogers and Castro, 1981) and this is likely to influence differentials in fertility between migrants and non-migrants (because migrants are concentrated in key childbearing ages). Use of such covariates and the range of migration measures which are possible will assist with the identification of any migration-fertility timing effect which may be playing out among recent migrants. The third aspect is the degree to which there is an elevated level of fertility among migrant groups once the other characteristics of the migrant have been taken into account (i.e. underlying differences in the age structure, differences by country of birth). Andersson (2004) used a relatively limited range of predictors because of incompleteness of data and presented the study as more of a demographic one. However, given the strong associations between migration and specific characteristics (notably age), the present analysis wishes to identify the extent to which any differential fertility among recent migrants may be related to the socio-demographic characteristics of the migrant group. Therefore, a fuller range of factors which are strongly related to childbearing will be included in the analysis.

Given the way in which new migrants to England and Wales are recorded by the NHSCR and enter the ONS LS this chapter will select a group of migrants who entered England and Wales in the 1991-2001 period for the first time and compare the fertility of this group in the 24 months after the 2001 census to continuously resident consistent cases resident throughout the 1991-2001 time period. Through selecting these two groups it will be possible to compare the relative risk of births for the two groups in the 24 months after the 2001 census. It is important to note that migrants in the samples selected include those moving from Northern Ireland and Scotland to England and Wales. The first 24 months after the 2001 census are used for two main reasons. Firstly, there is a higher degree of certainty that the LS members are resident in this period compared with later on in the decade. As discussed in Chapter 4, there is attrition in the LS, particularly among foreign-born female LS members. Therefore, to ensure that the results arrived at are as robust as possible it is only feasible to estimate the risk of birth for the period immediately after the 2001 census. Secondly, the 2001 census provides the full range of socio-economic variables from the census which may be important to consider in accounting for the fertility of migrants to England and Wales. For migrants arriving in years after the 2001 census; data is limited to that collected at the time of any birth.

8.3 Research questions

This chapter is concerned with one overarching question – what is the fertility of recent migrants to England and Wales relative to non-migrants in the 2001-2003 period? However, there are three specific questions on the fertility of migrants relative to non-migrants, which are crucial for answering this question:

1. Do migrants from the 1991-2001 period show a higher level of fertility in the period after the 2001 census compared to non-migrants?
2. Does the composition (socio-economic / demographic characteristics) of the migrant group lead to elevated fertility in comparison with non-migrants?
3. Is there an elevated level of fertility associated with the migration event itself (i.e. is there a higher likelihood of a birth to migrants who recently moved to England and Wales just before the 2001 census)?

Each of these three questions is interrelated and satisfactorily answering all three will provide a clearer insight on the fertility of recent migrants in England and Wales.

To answer each of the research questions outlined in section 8.3 two methods will be applied – a life-table of births to migrants and non-migrants by duration from the 2001 census and the use of a series of discrete-time hazards models. The two methods are similar in that they estimate the hazard of a birth in the 24 months following the 2001 census but the regression models do not use a continuous measure of time, control for the type of migrant (i.e. date of migration) and (in most cases) socio-economic characteristics which are collected at the 2001 census. Meanwhile, the hazard function from life table analysis does not control for the LS members characteristics and uses a continuous measure of time.

8.4 Method for life table analysis

As part of answering question one a life table will be specified with the hazard function and survivor function calculated. This will provide results on hazard and survival probabilities for migrants and non-migrants in the 24 months after the 2001 census.

8.4.1 Identifying migrant and non-migrant groups

Two broad groups will be selected for analysis:

- ***'Migrants'*** *– defined as LS members who entered the LS for the first time through a registration with a GP in the 1991-2001 period (NHSCR registration) OR who were identified at the 2001 census for the first time.*
- ***'Non-migrants'*** *– previously termed 'continually resident consistent cases' or 'Type 1 cases' in Chapter 4. These are persons who were continuously resident between 1991 and 2001 with no recorded embarkation or re-entry; all evidence suggests that these persons remained resident. (It is possible that these persons were born overseas and migrated to England and Wales at some point in the past and became part of the ONS LS). The key point about this group is that they did not migrate in the 1991-2001 period and are a 'stable' comparator.*

At a macro level the following criteria have been applied to the data to select LS members for all the analysis in this chapter:

- resident at the 2001 census (necessary for the covariates required and the start of the period of observation);
- did not enter the LS with a date of birth discrepancy (because there is the potential that such entries are erroneous and subsequent births will not be recorded);
- was traced at some point in time on the NHSCR (so births to these LS members are recorded).

8.4.2 Life table of births to migrants and non-migrants

Results will allow the identification of any difference in fertility between the migrant and non-migrant groups. The hazard function shows the conditional probability that an individual will experience a birth given that they have not experienced a birth in an earlier time period. Meanwhile, the survivor function cumulates the period-by-period risks of birth occurrence to assess the probability that an individual will not experience a birth.

Hazard function $\hat{h}(t_j) = \frac{n\ events_j}{n\ at\ risk_j}$

Survivor function $\hat{S}(t_j) = \frac{n\ who\ have\ not\ experienced\ the\ event\ by\ the\ end\ of\ time\ period\ j}{n\ in\ the\ data\ set}$

From Singer and Willett (2003).

8.5 Method for discrete-time logistic regression analysis

8.5.1 Person-period dataset for May 2001- April 2003

For the life table analysis and discrete-time hazards model analysis a person-period version of the ONS LS will be constructed. For each month of exposure to risk of birth a row will be created which includes each LS member at the 2001 census in the migrant or non-migrant group. The row will include the measures of the migration event (discussed shortly) for each LS member and whether or not there was a birth to the LS member at that duration. Each month in the May 2001-April 2003 period must be calculated and then the values for each of these months stacked in a woman-month observation format. Figure 8.1 illustrates the structure of the ONS LS dataset in the stacked person period format. The dataset is composed of the identification number for the LS member, the duration from the assumed date of migration to England and Wales to the 2001 census, whether or not there was a birth in the month, and whether or not there was a departure (emigration or death) in that month.

Figure 8.1: Example structure of person-period data from LS – duration from assumed migration date

CORENO (ID)	DURATION_SINCE _MIGRATION	BIRTH	MTHS_SINCE _2001_CENSUS	DEPARTURE	DOBYR	Covariate 1	Covariate 2..
1	15	0	1	0	1978		
1	16	0	2	0	1978		
1	17	0	3	0	1978		
1	18	0	4	0	1978		
1	19	0	5	0	1978		
1	20	0	6	0	1978		
1	21	0	7	0	1978		
1	22	0	8	0	1978		
1	23	0	9	0	1978		
1	24	0	10	0	1978		
1	25	0	11	0	1978		
1	26	0	12	0	1978		
1	27	0	13	0	1978		
1	28	0	14	0	1978		
1	29	0	15	0	1978		
1	30	0	16	0	1978		
1	31	0	17	0	1978		
1	32	0	18	0	1978		
1	33	0	19	0	1978		
1	34	0	20	0	1978		
1	35	0	21	0	1978		
1	36	0	22	0	1978		
1	37	0	23	0	1978		
1	38	1	24	0	1978		

8.5.2 Analytical strategy

In line with the three research questions outlined, a range of discrete-time logistic regression hazards models will be specified (Allison, 1982). It is necessary to specify different models because of the different measures of the migration process and also the aim to identify between compositional elements of the migrant sample (e.g. age, country of birth and other covariates) compared to the non-migrant group.

To answer each of the three research questions as defined previously it is necessary to specify models which progressively introduce the different measures and controls.

1. The first research question relates to identification of any kind of increased fertility regardless of the characteristics of the migrants / non-migrants and will be answered using life table analysis. Migrant groupings will be based on the use of the NHSCR registration date and migration indicator at the 2001 census.
2. The second research question is concerned with identification of any increased fertility in relation to the composition of the migrant group.

Here the characteristics of the migrant may be playing a part in any elevated level of fertility which is identified in 1. Therefore a range of covariates will be added into the models to control for common characteristics associated with fertility. These will include age, parity as of the 2001 census, education level, marital status, economic position and country of birth. In addition, interactions will be considered between age and parity and age and education level. The theoretical reason for the inclusion of these variables and the interaction terms is explained shortly.

3. The third question is concerned with different measures of the date of arrival of the migrant in England and Wales (based on earlier analysis in Chapter 6). Once the underlying characteristics of the migrant have been controlled for by inclusion of the range of socio-economic variables in question 2, is there any elevated fertility using the measures of migration available? To estimate the migration process a range of measures will be used based around the NHSCR date and whether the LS member was overseas 12 months before the 2001 census. These migrant groupings will be discussed shortly. In order to assess this the models control for the range of socio-economic characteristics as included in models specified in response to question 2 and use measures of the date of migration to England and Wales. The analysis in Chapter 6 identified that there is an association between the date of NHSCR registration and subsequent birth. Therefore, the date of NHSCR registration can only be used as one of a range of measures of the date of true migration. The migration event is measured using the NHSCR date and the migration indicator from the 2001 census which asked if the LS member was living overseas 12 months before.

A range of measures are used for the migration process and the characteristics of the migrant and non-migrant, these are discussed in the next section in relation to childbearing.

8.5.3 Measures – variables and substantive background to their inclusion

As already discussed, the outcome variable for this work will be whether there is a birth in the 24 months after the 2001 census.

Outcome variable – Birth in the 24 months after the 2001 census (May 2001 – April 2003). As already discussed, the outcome which the models seek to predict is whether there is a birth or not in the 24 months following the 2001 census. Exposure to risk of birth is granted to LS members at the 2001 census until a (any parity) birth, death or recorded NHSCR embarkation from the LS.

For each of the models a different set of variables will be included. As discussed in section 8.2 on the rationale and with reference to the second and third research questions, a set of control variables need to be specified. These variables are those which are the most theoretically important in accounting for fertility. The reason for the inclusion of each variable is explained in turn:

Unit of time – will be months since the 2001 census. Given the analysis is concerned with the timing of fertility after the 2001 census this variable is included to show the quarter of the 24 months after the 2001 census in which any birth occurred. This will differ from the life table analysis which uses a continued measure of time. It is necessary to use a grouped variable because of the relatively small numbers which fall into each of the 24 months whilst at the same time showing any variation in fertility through the time period. The use of six month groupings has also been informed by outputs from the life table analysis presented in the results section.

Age group – this variable uses five year age groups as of the 2001 census. It would be anticipated that among migrants births would be focussed among the younger age groups (20-25 years). For the non-migrant group a slightly later profile of fertility would be expected compared to migrants, in line with the prevailing age-specific fertility profiles for England and Wales. Tromans et al. (2008) provides a background indication on what might be expected for the migrant groups.

Parity as of the 2001 census – parity gives an indication of the number of children already born to the LS member. There is a prevailing two-child preference in England and Wales (Sigle-Rushton, 2008) with no one child preference (Jefferies, 2001). Therefore, high rates of progression from one to two children would be anticipated. Differences between migrants and non-migrants may show a more likely progression from parity zero to one if there has been some form of postponement of fertility before migration.

Age and parity interaction – the purpose is to identify the degree to which age and parity act together in fertility patterning. There are two aspects to this within the life course. The first is younger, earlier fertility which is related to a high fecundity and limited life chances (Arai, 2007). The second is recuperation as a result of postponement at older ages which leads to lower parity births at higher ages (Berrington, 2004).

Education level – Timing of childbearing is related to time enrolled in education and subsequent post-education postponement (Ní Bhrolcháin and Beaujouan, 2012; Rendall et al., 2005) which extends transitions to adulthood (Blossfeld and Huinink, 1991).

Age and education level interaction – the effects of education on the likelihood of a birth change with age. Using age and education as main effects alone would obscure any relationship between the two variables which may be playing out. For example, among younger women, those with no or limited qualifications typically show a higher likelihood birth and to do older, highly qualified women who have a higher likelihood of giving birth. Rendall and Smallwood (2003) used the ONS LS to identify the associations between age and education in childbearing. The average age of entry to motherhood was found to be five years later for women with higher qualifications than for those without.

Economic position – fertility is related to labour market attachment and position within the labour market (in terms of the degree of attachment, level of employment etc) and also related to educational attainment. Kneale and Joshi (2008) identified postponement across cohorts and impact on eventual childlessness. Differential economic position is likely to arise from the country of origin with Pakistani and Bangladeshi females likely to be more economically disadvantaged (Dustmann and Fabbri, 2005).

Marital status – There is a continued rise in the number of births outside of marriage, cohabiting women still have an overall level of fertility below that of married women but above that of other unmarried women (O'Leary et al., 2010). Among the migrant groups it is possible that marital status will be of greater importance in predicting fertility because of differing preferences / norms among the migrant groups (Peach, 2006).

Country of birth – given this research is primarily concerned with the fertility of recent migrants the country of origin must be considered. For example, it is likely that among Bangladeshi and Pakistani women fertility will be higher than for women from other countries (Coleman and Dubuc, 2010; Peach, 2006). Country of birth patterning may be related to the reasons for the migration. For example, greater marriage and family formation among Bangladeshi and Pakistani groups compared to East African groups may lead to differing likelihood of a birth in the period of observation. Rendall and Salt (2005) discussed characteristics of the foreign-born population at the 2001 census and identified that for immigrants from higher-income countries there was a lower likelihood of remaining permanently resident in the UK.

High and low income country of birth groupings – Given findings made by Rendall and Ball (2004) that emigration higher among migrants arriving from higher income countries (a higher level of short-term migration) a variable based on the country-groupings used may be included in the models specified. It is possible that women from lower income countries will show a higher risk of a birth in the period of observation because they are more likely to remain resident (i.e. not lost to follow-up) and may also have migrated to England and Wales for partnership and family building reasons (Peach, 2006) compared to short term employment-related migration among migrants originating from higher income countries.

Note that country of birth and the Rendall and Ball country groupings will be tested in repeated versions of the same model to check on the insights they provide for fertility by country of origin. It will be necessary to choose between these two options for the final models specified.

Migration – a range of variables will be used to measure the date of migration to England and Wales, these will use the date of NHSCR registration (with a range of sensitivities applied) and the 2001 census question on place of residence (overseas) 12 months before. The full details on the measures to be used are discussed in an upcoming subsection.

8.5.4 Model selection

The size of the sample for this analysis (2.3 million person months of exposure in a person-period format), means it is likely that most variables included as covariates will be statistically significant. Indeed, even very small substantive effects of the covariates included in the models may be statistically significant because of the sample size. Model selection will be carried out using an automated forward stepwise function to determine variables which should be included to lead to the most statistically parsimonious model. However, even with an automated procedure (Likelihood Ratio is used in this case), the large sample size may mean that it is highly likely that many of the main effects and interaction terms specified are included. Forward selection starts with the independent variable which is the best predictor of the dependent variable, checks that the coefficient is significantly different from zero at the 5% level and then progressively adds other variables which improve the prediction the most (Dugard et al., 2010). Variables in the model are checked for their statistical significance after the inclusion of the new variable and if they are no longer significant are removed from the model. The process continues until no more variables pass the criterion.

A key point to be noted is that final variables and interaction terms selected for inclusion will be based on those which are retained by the forward stepwise function for the continuously-resident non-migrant group which is a larger sample than for the migrant group. It is necessary to retain / base the selection of covariates from the automated procedure on the non-migrant group because the analysis will compare the effect of key variables across the migrant and non-migrant groups in different hazard models with the alternative measures of migration. In the final model some groupings for key control variables are collapsed because of small sample sizes.

The formulae for discrete time logistic regression hazards models specified is detailed below. Hazard of a birth ($logit(h_{ti})$) is the outcome. In this case, time (t) is measured in 6 month blocks (as discussed in the measures section) and therefore assumes a constant hazard over each block of time. Covariates are fixed in this case (from the 2001 census) but with a value for each time point (x_{ti}) and time is duration from the 2001 census ($a(t)$).

$$logit(h_{ti}) = log\left(\frac{h_{ti}}{1-h_{ti}}\right) = a(t) + \beta' x_{ti}$$

(Steele, 2005).

8.5.5 Comparing fertility across groups of migrants and non-migrants

Identification of the date of migration from the NHSCR date of registration seems to be problematic because of the link between registration and subsequent fertility which has been discussed previously. Therefore, the analysis uses a range of measures of the migration process. Assumptions are made about the true, unrecorded date of migration for new migrants into the ONS LS to create additional variables on the date of migration to England and Wales. The migrant groups identified in Chapters 6 and 7 are utilised. Again, as already stated earlier in the chapter, it must be noted that migrants in the samples selected include those moving from Northern Ireland and Scotland to England and Wales.

1. Continuous measures of date of migration

Using information available for LS members who entered England and Wales and the ONS LS between 1991 and 2001 a range of sensitivity analyses of the date of migration from the NHSCR is made. The NHSCR date itself is used along with three estimated dates of the earliest and latest dates of migration and a mid-range estimate between the two. These values are established using all the information on the whereabouts of the LS member. In this way the process is similar to that used in Chapter 4 to identify the residence trajectories for LS members. This process could be described as 'case by case' in that the individual level LS member information is being used rather than a 'source by source' approach concerned with the source of the data. The alternative estimates of the duration from migration to the 2001 census using the range of sensitivities are then used to estimate the fertility of migrants entering the ONS LS between 1991 and 2001 relative to LS members who were continuously resident.

With the period of observation for this sample starting from the 2001 census (April 2001), the important step is to establish the duration from entry to the LS as of the 2001 census. Looking at each LS member in turn it is possible to identify what the date of migration to England and Wales may have been. This is different to the date of entry to the ONS LS (see explanation in Chapter 6). For migrants four assumptions on the date of migration to England and Wales will be made as follows.

Migration date 1 uses the date of NHSCR registration. The duration from the date of NHSCR registration to the 2001 census is then used in analysis.

Migration date 2 uses the maximum duration identifying the earliest date in the 1991-2001 period at which the migration could have occurred. For some LS members where there is no recorded entry before the 2001 census this date will be the 1991 census. For other LS members there is a recorded date registered on the NHSCR, but this may not be the actual date of arrival. For example, an LS member registers with a GP in 1994, and hence there is a recorded date of entry on the NHSCR as 1994. However, the LS member may have arrived as early as April 1991 (immediately after the 1991 census) but not registered with a GP until a later date. Therefore the maximum duration must be from April 1991.

Migration date 3 is the minimum duration identifying the latest date in the 1991-2001 period at which the migration could have occurred. For some LS members where there is no recorded entry before the 2001 census this date will be the 2001 census. For example, an LS member indicates that they were living overseas 12 months before the 2001 census and there is no date of entry to England and Wales so the minimum duration would be counted from April 2001.

Migration date 4 is a mid-range duration estimate, calculated by identifying the mid-point between the maximum and minimum. This is arrived at by subtracting the maximum (Migration date 2) from the minimum (Migration date 3).

Under the different migration date assumptions the duration since arrival in England and Wales variable used in the regression analysis will change. The duration (in months) from migration to the 2001 census will be calculated. The different estimates of duration from migration will be used to estimate the fertility of migrants who entered the ONS LS between 1991 and 2001 relative to LS members who were continuously resident.

2. Measures of date of migration using the 2001 census variable on place of residence 12 months before

Using information from the 2001 census on if the LS member is living overseas and the NHSCR date of registration with a GP a range of migrant groups are constructed. Table 8.1 presents a summary of the different migration groups which will be used before Table 8.2 elaborates on models to be specified.

Table 8.1: Migration groupings to be used

Migration grouping	Migration variable used	Purpose
Detailed migrant groupings from 2001 census	Uses date of NHSCR registration and variable on if overseas 12 months before	Uses all the available migration data from the NHSCR and 2001 census to construct variables for the different migrant groups at the 2001 census.
Variable on if overseas 12 months before 2001 census	Variable from 2001 census on if overseas 12 months before.	Does not use the NHSCR data and is a separate / self-reported measure of the date of migration in relation to the 2001 census. Also to be used in a separate model (see below).
Duration from migration to 2001 census using range of estimates	Uses date of NHSCR registration.	A range of estimated date of arrival (as described above).
Migrant groupings from 2001 census including non-recent migrants (1971-1991)	(Alternative migrant groupings). Uses date of NHSCR registration and variable on if overseas 12 months before. Also splits the non-migrant group to use the variable on if ever a migrant to England and Wales.	Seeks to identify if the comparator group should be split down and any difference within the continually resident group depending on if a migrant in the past (i.e. a higher fertility among non-recent migrants).
Only LS members overseas 12 months before the 2001 census	-	Characteristics associated with a birth just for recent migrants.

Table 8.2: Discrete-time hazard models to be specified

Model	Included	Relevance to research questions	Migration variable used	Purpose
Model 1	*Migrants / non-migrants*	1 (migration), 2 (migration and age) and 3 (migration and covariates).	NHSCR and 12 month indicator at 2001 census	Fertility of migrants compared to non-migrants for duration from 2001 census. Allows comparison to life table results.
Model 2	*Migrants / non-migrants*	2 and 3.	NHSCR and 12 month indicator at 2001 census	No duration from migration variable. Comparing the effect of covariates across migrant and non-migrant groups
Model 3	*Migrants*	2 and 3.	12 month indicator at 2001 census	Using a dummy variable (derived from 'MIGP0') on if the LS member was overseas 12 months before the 2001 census. Independent of NHSCR entry information.
Model 4	*Migrants*	2 and 3.	NHSCR (and range of sensitivities)	Using the dates of migration from section 8.5.5 to estimate duration to the 2001 census (April 2001) which is the start point of the 24 month period where the LS member is under observation.
Model 5	*Migrants / non-migrants*	2 and 3.	NHSCR and 12 month indicator at 2001 census	Groups of migrants and non-migrants selected depending on date of NHSCR registration and whether overseas 12 months before the 2001 census. Categories combined from Model 1. Also identifies non-recent migrants.
Model 6	*Migrants*	2 and 3.	(Just sample overseas) 12 month indicator at 2001 census	Selects sample of migrants overseas 12 months before the 2001 census to analyse fertility

8.6 Results from descriptive statistics for migrants and non-migrants

Two groups of LS member will be compared in this analysis; these will consist of migrants who arrived in England and Wales between 1991 and 2001 and persons who were continuously resident consistent cases between 1991 and 2001. In this section additional background information is provided on the sample for analysis.

8.6.1 Descriptive statistics of variables to be included in the models

For each of the variables to be included in the modelling it is necessary to consider the percentage distribution across the variable and cell count sizes, for the migrant and non-migrant groups. Implications of cell counts for the modelling are important and so each variable is disaggregated for the migrant and non-migrant group. In this section frequencies are shown for the number of LS members in each of the migrant and non-migrant groups in month 1 from the person-period dataset.

Each variable is explained along with the counts for this variable for each of the four dates of migration being used. With migration date 1 there are fewer cases than for migration dates 2-4 because these include persons who arrived at the 2001 census for the first time and did not register with a GP before this date. For migration dates 2-4 the LS member does not have to have registered with a GP because through using everyone at the census and the information on their migration and residence in the 1991-2001 period it is possible to attribute the earliest possible and latest possible date of migration to England and Wales (and using these estimate a mid point date when the LS member may have moved).

Month *– this variable identifies the month in the 24 months from May 2001 to April 2003. Where the LS member was resident and had not given birth then they are included in the sample. Once a birth or death / embarkation from the NHSCR occurred the person is removed from the sample. (i.e. someone who gives birth in month 12 is exposed to risk for the first 12 months in the 24 month period).*

***Duration from migration at 2001 census** – this variable calculates the number of months from the date of migration under each assumption to the month of the 2001 census – April 2001.*

***Grouped age** at 2001 census. LS members under the age of 15 and over the age of 45 as of the 2001 census have been removed from the sample.*

Table 8.3 shows the counts for each of the age groups used in the analysis. These show that overall, there is a more even spread in the age for non-migrants relative to the migrant group. Among the migrant group there is a greater concentration of the sample in ages under 35 years at the 2001 census. Indeed, just over 20% of the sample in the migrant group was older than 35 years at the 2001 census compared to 38% for the non-migrant group. In Table 8.4 the same figures are shown for the other migration assumptions. There is a similar profile in the age distribution with slightly higher percentages in the older ages.

Table 8.3: Grouped age variable for non-migrants and migrants (Migration date 1) in the 1991-2001 period and continually resident consistent cases 1991-2001

	Non-migrants		Migrants		Total	
	N	%	N	%	N	%
15-19	13,818	15.0	481	8.7	14,299	14.6
20-24	12,780	13.8	1,155	20.9	13,935	14.2
25-29	14,011	15.2	1,567	28.3	15,578	15.9
30-34	16,530	17.9	1,161	21.0	17,691	18.1
35-39	18,232	19.7	748	13.5	18,980	19.4
40-44	16,952	18.4	416	7.5	17,368	17.7
Total	92,323	100	5,528	100	97,851	100

Own elaboration based on ONS LS, September 2011.

Table 8.4: Grouped age variable for non-migrants and migrants (Migration date 2, 3 & 4) in the 1991-2001 period and continually resident consistent cases 1991-2001

	Non-migrants		Migrants		Total	
	N	%	N	%	N	%
15-19	13,818	15.0	570	8.9	14,388	14.6
20-24	12,780	13.8	1,260	19.6	14,040	14.2
25-29	14,011	15.2	1,727	26.9	15,738	15.9
30-34	16,530	17.9	1,410	22.0	17,940	18.2
35-39	18,232	19.7	908	14.1	19,140	19.4
40-44	16,952	18.4	546	8.5	17,498	17.7
Total	92,323	100	6,421	100	98,744	100

Own elaboration based on ONS LS, September 2011.

***Country of birth** – using the 'COBP0' variable country of birth groupings for LS members at the 2001 census can be created. These are comparable to the ONS FM1 'Families and fertility' series. Note that persons continually resident in the 1991-2001 period may have been born overseas and migrated to England and Wales before the 1991 census.*

Country of birth information for each LS member is shown in Table 8.5. In interpreting this it is important to recall that the non-migrant sample includes persons who were born overseas and migrated to England and Wales at some point before the 1991 census. The country of birth groupings used corresponds with those used by the ONS in the Families and Fertility 'FM1' volume. Among the migrant group there is a large sample of LS members from the European Union and also 'uncoded'. Migrants from the UK are included in this sample because the LS migrants include those persons who come from Scotland and Northern Ireland. Table 8.6 shows the same information for the samples from the other migration assumptions. These assumptions lead to a higher percentage of UK born women and relatively modest changes for the other countries of birth.

Table 8.5: Country of birth variable for non-migrants and migrants (Migration date 1) in the 1991-2001 period and continually resident consistent cases 1991-2001

	Non-migrants		Migrants		Total	
	N	%	N	%	N	%
UK	86,480	93.7	502	9.1	86,982	88.9
Ireland	321	0.3	152	2.7	473	0.5
Australia, Canada, New Zealand	226	0.2	311	5.6	537	0.5
India	709	0.8	408	7.4	1,117	1.1
Pakistan	562	0.6	342	6.2	904	0.9
Bangladesh	397	0.4	194	3.5	591	0.6
East Africa	459	0.5	139	2.5	598	0.6
Other EU 25	657	0.7	815	14.7	1,472	1.5
Uncoded / other	2,512	2.7	2,665	48.2	5,177	5.3
Total	92,323	100	5,528	100	97,851	100

Own elaboration based on ONS LS, September 2011.

Table 8.6: Country of birth variable for non-migrants and migrants (Migration date 2, 3 & 4) in the 1991-2001 period and continually resident consistent cases 1991-2001

	Non-migrants		Migrants		Total	
	N	%	N	%	N	%
UK	86,480	93.7	992	15.4	87,472	88.6
Ireland	321	0.3	176	2.7	497	0.5
Australia, Canada, New Zealand	226	0.2	329	5.1	555	0.6
India	709	0.8	436	6.8	1,145	1.2
Pakistan	562	0.6	384	6.0	946	1.0
Bangladesh	397	0.4	217	3.4	614	0.6
East Africa	459	0.5	150	2.3	609	0.6
Other EU 25	657	0.7	875	13.6	1,532	1.6
Uncoded / other	2,512	2.7	2,862	44.6	5,374	5.4
Total	92,323	100	6,421	100	98,744	100

Own elaboration based on ONS LS, September 2011.

> ***Household marital status*** *– at the 2001 census the 'MHUTYP0' variable gave the minimal household type of LS member. Note that this variable includes the parity of the LS member at the census.*

As the description of this variable indicates, the household type and marital status of the LS member at the 2001 census includes birth parity. This is a 'portmanteau' variable as it consists of marital status and parity. Table 8.7 shows the spread of the samples for this variable. Among the migrant group there are more persons in the 'communal establishment / other / missing' category and the 'married couple only' category compared with the non-migrants. There are fewer migrants in the 'lone parent with dependent child' and 'cohabiting couple with dependent children' categories. When the other migration assumptions are used in Table 8.8 there are modest changes to the percentages with the sample generally becoming more similar to that for the non-migrant group.

Table 8.7: Household marital status variable for non-migrants and migrants (Migration date 1) in the 1991-2001 period and continually resident consistent cases 1991-2001

	Non-migrants		Migrants		Total	
	N	%	N	%	N	%
Person in communal est/ not recorded / other / missing / no value	1,177	1.3	415	7.5	1,592	1.6
Unmarried adult	21,758	23.6	1,394	25.2	23,152	23.7
Lone parent with dependent child	12,129	13.1	428	7.7	12,557	12.8
Married couple only	8,411	9.1	901	16.3	9,312	9.5
Cohabiting couple only	7,410	8.0	453	8.2	7,863	8.0
Married couple with dependent children	35,168	38.1	1,824	33.0	36,992	37.8
Cohabiting couple with dependent children	6,270	6.8	113	2.0	6,383	6.5
Total	92,323	100	5,528	100	97,851	100

Own elaboration based on ONS LS, September 2011.

Table 8.8: Household marital status variable for non-migrants and migrants (Migration date 2, 3 & 4) in the 1991-2001 period and continually resident consistent cases 1991-2001

	Non-migrants		Migrants		Total	
	N	%	N	%	N	%
Person in communal est/ not recorded / other / missing / no value	1,177	1.3	444	6.9	1,621	1.6
Unmarried adult	21,758	23.6	1,602	24.9	23,360	23.7
Lone parent with dependent child	12,129	13.1	554	8.6	12,683	12.8
Married couple only	8,411	9.1	997	15.5	9,408	9.5
Cohabiting couple only	7,410	8.0	529	8.2	7,939	8.0
Married couple with dependent children	35,168	38.1	2,143	33.4	37,311	37.8
Cohabiting couple with dependent children	6,270	6.8	152	2.4	6,422	6.5
Total	92,323	100	6,421	100	98,744	100

Own elaboration based on ONS LS, September 2011.

***Economic position** – grouped from the 'ECOP80' variable showing the type of economic activity the LS member was involved in as of the 2001 census.*

In some of the individual economic activity categories there were small numbers emerging which could be problematic for the regressions. Therefore the variable presented uses a grouped economic position variable. Table 8.9 shows the sample distribution for this variable. The migrant group is composed of far fewer women who work part time (9%) compared with the non-migrant group (22%). Interestingly, the full time figures are comparable. As would be anticipated with the migrant group, there are higher percentages for students with 20% of the migrants in this group compared to 11% of non-migrants. Table 8.10 again shows a generally similar profile to the migrants in Table 8.9 but with percentages that move towards the non-migrant sample.

Table 8.9: Economic position variable for non-migrants and migrants (Migration date 1) in the 1991-2001 period and continually resident consistent cases 1991-2001

	Non-migrants		Migrants		Total	
	N	%	N	%	N	%
Not recorded / other / no value	6,312	6.8	541	9.8	6,853	7.0
Employed Part time	20,091	21.8	468	8.5	20,559	21.0
Employed Full Time	34,327	37.2	1,684	30.5	36,011	36.8
Self employed	3,722	4.0	162	2.9	3,884	4.0
Seeking work / Retired / Permanently Sick	5,099	5.5	314	5.7	5,413	5.5
Student	10,059	10.9	1,121	20.3	11,180	11.4
Looking after home	12,713	13.8	1,238	22.4	13,951	14.3
Total	92,323	100	5,528	100	97,851	100

Own elaboration based on ONS LS, September 2011.

Table 8.10: Economic position variable for non-migrants and migrants (Migration date 2, 3 & 4) in the 1991-2001 period and continually resident consistent cases 1991-2001

	Non-migrants		Migrants		Total	
	N	%	N	%	N	%
Not recorded / other / no value	6,312	6.8	614	9.6	6,926	7.0
Employed Part time	20,091	21.8	586	9.1	20,677	20.9
Employed Full Time	34,327	37.2	2,015	31.4	36,342	36.8
Self employed	3,722	4.0	201	3.1	3,923	4.0
Seeking work / Retired / Permanently Sick	5,099	5.5	379	5.9	5,478	5.5
Student	10,059	10.9	1,229	19.1	11,288	11.4
Looking after home	12,713	13.8	1,397	21.8	14,110	14.3
Total	92,323	100	6,421	100	98,744	100

Own elaboration based on ONS LS, September 2011.

***Education level** – using the 'HLQP0' variable gives the highest qualification for the LS member. This was derived from qualifications and professional qualifications questions at the 2001 census.*

Education level information is simplified in the regression models presented in section 8.8. Here the full information on what each level means is given in Table 8.11 and 8.12. These show a higher percentage of LS members with no qualifications in the migrant category (21%) relative to the non-migrants (14%). This is important to note as the no qualifications coding is distinct from 'not recorded' where there are quite low percentages for migrants compared to the non-migrants. Overall, the migrant group is more concentrated among the higher qualifications with just under 40% of the sample in the level 4/5 category. This compares with just under 20% for the non-migrant group. At the lower end of the qualification scale there are fewer migrants and more non-migrants. For migration dates 2-4 there are similar percentages apart from the no qualifications group where around 20% of the migrant group are in this category and 14% of non-migrants. Table 8.13 allows comparison of the UK education groups used in the analysis against international levels (the International Standard Classification of Education (1997)).

Table 8.11: Education level variable for non-migrants and migrants (Migration date 1) in the 1991-2001 period and continually resident consistent cases 1991-2001

	Non-migrants		Migrants		Total	
	N	%	N	%	N	%
Not recorded / other / missing / no value	3,677	4.0	94	1.7	3,771	3.9
No academic or professional qualifications	12,873	13.9	1,131	20.5	14,004	14.3
Level 1: CSEs (grades 2-5), GCSEs (grades D-G), 1-4 CSEs (grade 1), 1-4 GCSEs (grades A-C)	20,594	22.3	390	7.1	20,984	21.4
Level 2: 5+O levels, 5+CSEs (grade1), 5+GCSEs (grades A-C) etc, 1 A level, 1-3 AS levels, NVQ level 2, Intermediate GNVQ	25,040	27.1	666	12.0	25,706	26.3
Level 3: 2+ A levels, 4+ AS levels, Higher Sc Cert, NVQ 3, Advanced GNVQ	10,406	11.3	776	14.0	11,182	11.4
Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	17,675	19.1	2,177	39.4	19,852	20.3
Other qual/ level unknown- Other qual (eg City and Guilds), Other Prof qual	2,058	2.2	294	5.3	2,352	2.4
Total	92,323	100	5,528	100	97,851	100

Own elaboration based on ONS LS, September 2011.

Table 8.12: Education level variable for non-migrants and migrants (Migration date 2, 3 & 4) in the 1991-2001 period and continually resident consistent cases 1991-2001

	Non-migrants		Migrants		Total	
	N	%	N	%	N	%
Not recorded / other / missing / no value	3,677	4.0	112	1.7	3,789	3.8
No academic or professional qualifications	12,873	13.9	1,303	20.3	14,176	14.4
Level 1: CSEs (grades 2-5), GCSEs (grades D-G), 1-4 CSEs (grade 1), 1-4 GCSEs (grades A-C)	20,594	22.3	496	7.7	21,090	21.4
Level 2: 5+O levels, 5+CSEs (grade1), 5+GCSEs (grades A-C) etc, 1 A level, 1-3 AS levels, NVQ level 2, Intermediate GNVQ	25,040	27.1	827	12.9	25,867	26.2
Level 3: 2+ A levels, 4+ AS levels, Higher Sc Cert, NVQ 3, Advanced GNVQ	10,406	11.3	882	13.7	11,288	11.4
Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	17,675	19.1	2,477	38.6	20,152	20.4
Other qual/ level unknown- Other qual (eg City and Guilds), Other Prof qual	2,058	2.2	324	5.0	2,382	2.4
Total	92,323	100	6,421	100	98,744	100

Own elaboration based on ONS LS, September 2011.

Table 8.13: Qualifications in the UK and their equivalent ISCED-97 levels

NVQ/SVQ combined (ONS)	ISCED-97 (OECD)
Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	ISCED 3C
Level 2: 5+O levels, 5+CSEs (grade1), 5+GCSEs (grades A-C)	ISCED 3C
Level 3: 2+ A levels, 4 + AS levels	ISCED 3A
Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	ISCED 5B

Source: Adapted from Schneider, S. L. (2008) The application of the ISCED-97 to the UK's educational qualifications. In Schneider, S. L. (ed.) The International Standard Classification of Education – An Evaluation of Content and Criterion Validity for 15 European Countries, The Mannheim Centre for European Social Research (MZES), Mannheim, pp. 281-300. (http://www.mzes.uni-mannheim.de/publications/misc/isced_97/schn08e_the_application_of_the_isced-97_to_the_uks_educat.pdf).

8.6.2 Migrant status before the 2001 census – sample information

Overseas 12 months prior to the 2001 census *– using the 'MIGP0' allows identification of those LS members who were overseas 12 months before the 2001 census.*

As noted in the opening part of this subsection there is a slight difference in the sample size between the different migration dated samples used. In this subsection there is a brief explanation of the number of LS members at the 2001 census and in the migrant and non-migrant groups and overseas 12 months before the 2001 census. Frequencies here are again for LS members' resident in month 1 of the 24 months selected for analysis.

Table 8.14 shows the sample of LS members who were overseas 12 months before the 2001 census for the migrants under migration date 1 and migration dates 2-4. In both cases 89% of the sample was not overseas 12 months before the census.

Table 8.14: Migration indicator for 12 months before the 2001 census for migrants (Migration date 1) in the 1991-2001 period

Group	Migration date 1	Migration dates 2-4
	N	N
Not overseas 12 months before 2001 census	4,920	5,762
Overseas 12 months before 2001 census	608	659
Total	5,528	6,421

Own elaboration based on ONS LS, September 2011.

Breaking down the sample of those overseas gives an indication of those LS members who were not recorded as entering the LS in the 1991-2001 period and arrived at the 2001 census. In Table 8.14 this is shown as being 51 extra cases for migration dates 2-4 (659-608). For migration dates 2-4 there are 842 more individuals than under migration date 1 who were not overseas 12 months before the 2001 census, did not register with a GP in the 1991-2001 period and were at the 2001 census.

The above covariates will be used in section 8.8 to estimate the risk of birth in the May 2001-April 2003 period for migrants, non-migrants and the whole sample of LS members. The next section presents results for a life table of births to migrants and non-migrants.

8.7 Results from life table estimates of birth hazards for migrants and non-migrants: May 2001 – April 2003

In this section a life table for births to migrants and non-migrants in the 24 months following the 2001 census are presented. The pattern of coefficients for duration since April 2001 in the model should be similar to the hazard functions presented in this section.

8.7.1 Life tables for births to migrants and non-migrants

Table 8.15 presents the relative risk of birth for migrants and Table 8.16 presents the same analysis for non-migrants. Migration date 1 is being used in this analysis and the sample is composed of women who were aged 15-44 at the 2001 census. In both tables grouped months have been used because of small numbers of departures (deaths and embarkations) which cannot be published (due to ONS statistical disclosure controls). These tables show that migrants who entered the LS in the 1991-2001 period had a substantially higher likelihood of giving birth in the 24 months after the 2001 census. In both cases the number of births to migrants and the comparator group are similar for each month in the period. As would be expected given the findings on departures from the ONS LS among migrant groups in Chapter 4, there is a higher number of recorded embarkations for migrants (shown as 'departures' which includes deaths) relative to non-migrants. Births in month 0 are taken into account in calculating women exposed to risk in month 1.

Table 8.15: Life table showing births in the May 2001-April 2003 period to non-migrants (migrant date 1)

Months since census	Women exposed to risk of birth	Births	Departures (deaths and embarkations)	Hazard function - proportion of women giving birth in the month	Survivor function - proportion of all women still without a birth
0	92,326	369	-	-	1
1-3	274,735	1,109	15	0.012	0.988
4-6	271,448	1,193	21	0.013	0.975
7-9	268,057	1,039	29	0.012	0.964
10-12	265,137	994	35	0.011	0.953
13-15	262,056	1,121	24	0.013	0.940
16-18	258,790	1,100	28	0.013	0.928
19-21	255,651	1,037	30	0.012	0.917
22-24	252,685	1,026	27	0.012	0.906

Own elaboration based on ONS LS, September 2011.

Table 8.16: Life table showing births in the May 2001-April 2003 period to migrants (migrant date 1)

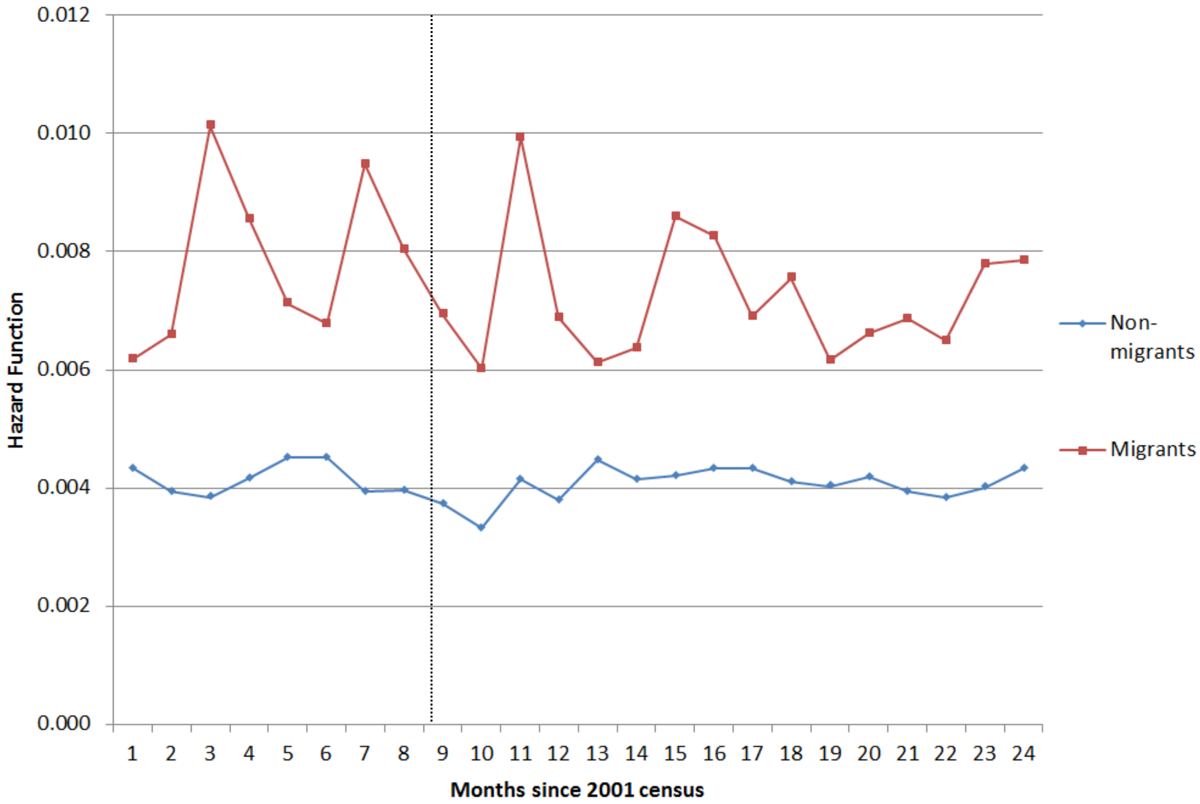
Months since census	Women exposed to risk of birth	Births	Departures (deaths and embarkations)	Hazard function - proportion of women giving birth in the month	Survivor function - proportion of all women still without a birth
0	5,529	33	-	-	1
1-3	16,391	125	6	0.023	0.977
4-6	16,018	120	14	0.022	0.955
7-9	15,692	128	10	0.024	0.932
10-12	15,376	117	18	0.023	0.911
13-15	15,081	106	10	0.021	0.891
16-18	14,789	112	16	0.023	0.871
19-21	14,508	95	11	0.020	0.854
22-24	14,245	105	6	0.022	0.835

Own elaboration based on ONS LS, September 2011.

Figure 8.2 presents the hazard function (risk of birth) for women in the migrant and non-migrant groups. This shows that throughout the 24 month period there is a substantially higher risk of birth among the migrant group. Although in the first 9 months the migrant group fertility could be associated with registration with a GP at the time of conception / pregnancy (denoted by the dashed line), the level of hazard is above that of the non-migrants for the whole time series of observation. Compared to the non-migrant group there is a greater fluctuation in the hazard for the migrant group. Among the non-migrant group the hazard level remains more consistent around 0.004 for the whole of the 24 month period. In month 10 (February 2002) there is a slightly lower hazard, but this is the only real drop in the series for the non-migrants.

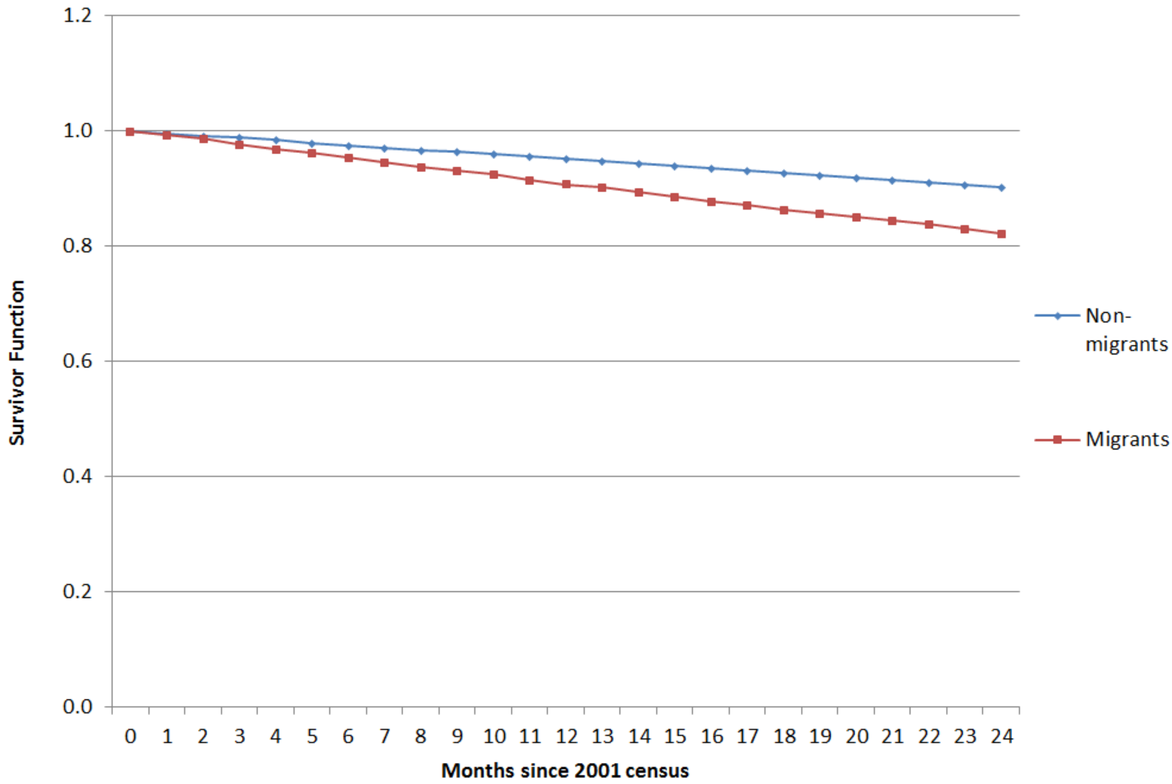
In Figure 8.3 the cumulative risk is used to calculate the survivor function which shows the proportion of women at each month in the period of observation who have not given birth. This provides information on the cumulative hazard for each group. For migrants the distribution function at the end of the period is half that for the non-migrants. The migrant group is almost twice as likely to give birth compared to the non-migrant group. Across the time period there is little change in the slope of the lines.

Figure 8.2: Estimated hazard functions for birth in the 24 months after the 2001 census (May 2001-April 2003) for migrants (NHSCR registration 1991-2001) and non-migrants (1991-2001)



Own elaboration based on ONS LS, September 2011.

Figure 8.3: Estimated survivor functions for birth in the 24 months after the 2001 census (May 2001-April 2003) for migrants (NHSCR registration 1991-2001) and non-migrants (1991-2001)



Own elaboration based on ONS LS, September 2011.

8.7.2 Summary

This section has shown that among the migrant group who entered the LS at some point in the 1991-2001 period there is a higher risk of birth in the period immediately after the 2001 census. Therefore, answering research question one, relative to non-migrants, migrants have a higher risk of birth. At the end of the 24 month period migrants are twice as likely to have given birth compared to non-migrants. However, the values arrived at in this analysis do not take into account the socio-demographic characteristics of the migrants relative to the non-migrants.

The next section is concerned with using a series of discrete-time hazards models to estimate the risk of a birth controlling for socio-economic characteristics of the migrant (as outlined in research question two) and identifying if there is an association between the date of arrival in England and Wales and subsequent birth (as outlined in research question three).

8.8 Results from hazards models of birth in the May 2001- April 2003 period

As outlined in the methods section, both a life table of births to the migrant and comparison group and discrete-time hazards models are needed in this analysis. In Section 8.7 it was shown that migrants to England and Wales in the 1991-2001 period had a higher risk of giving birth in the 24 months after the 2001 census. However, as outlined in section 8.2 which discussed the rationale for this work and the characteristics of the migrant sample, the life table analysis does not take into account the characteristics of the migrant group which may be leading to differential fertility compared to non-migrants.

In this section, research questions two and three which are concerned with the identification of a higher level of fertility among the migrant group which may be arising from the characteristics of that group and any timing effect among recent migrants are approached. The series of discrete-time hazard regression models presented in this section allows the identification of the impact of the socio-demographic characteristics of the migrant on fertility to be identified statistically and also through using a range of measures of migration any discernible elevated fertility after the migration event may be identified. It is necessary to use a range of measures of the migration process (as outlined in the methodology section) because each measure reflects the migration process differently.

We note that month, age and parity interactions, education level, economic position and country of birth are not statistically significant across all models but retain these variables for consistency across all the models and the range of measures of the migration process which have been used in the chapter. Indeed, the forward stepwise regression function retained the variables included for the non-migrant group and to ensure comparability between the migrant and non-migrant groups these have been retained across the models presented in the section. This is also the case for the interaction terms, many of which are not statistically significant, but which were retained for the non-migrant group and are therefore retained for the migrant group. Repetition of the same analysis for all models without the non-significant interaction terms did not change the values of remaining covariates.

Across the range of models, in all except the final model using the sample of LS members overseas 12 months before the 2001 census, the Rendall and Ball country groupings were not used. This is because there was not a statistically significant difference in the coefficients between the groupings whereas inclusion of the country of birth variable allowed identification of differing risk of birth for specific countries (e.g. Pakistan, Bangladesh (as hypothesised earlier) and Ireland).

8.8.1 Results from Model 1 – comparing the risk of birth for migrants and non-migrants

As an evolution of the life table results presented earlier, Table 8.17 presents a series of regression models for different types of migrant and progressively builds in covariate terms in order to appreciate the differences in the model when these are included.

Model 1A just shows just the two years since the 2001 census cut into quarters and the migrant groups. This shows the relative level of fertility for the different migrant groups for the two years after the census without controlling for the other socio-economic variables and interaction terms which can be included. Across the different months from April 2001 there is no substantial change in the likelihood of birth although in the 19-24 month period there is a drop in the odds ratio compared to months 1-6 (the reference category). The results from this model show that relative to those LS members who were not a migrant 1971-1991 (and continuously resident consistent cases 1991-2001) the migrants (1971-1991) had virtually no difference in their fertility after the 2001 census with similar odds ratios. In contrast, migrants who entered the ONS LS through an NHSCR registration (1991-2000) who were not overseas 12 months before the 2001 census had a higher level of fertility in the 24 months after the 2001 census with an odds ratio of 1.8. Those migrants who entered between 1991 and 2000 and were overseas, 12 months before the 2001 census have only a slightly elevated likelihood of giving birth in the

24 months after the 2001 census (odds ratio of 1.2 but not statistically significant in comparison to the reference category). Matching the findings in Chapter 6 and 7 with regard to the late registration of migrant LS members when they conceive / fall pregnant, those migrants with an NHSCR entry between April 2000 and April 2001 who were not actually overseas in April 2000 have the highest likelihood of a birth (odds ratio of 2.5). In contrast, the group of 'genuine' recent migrants who said that they were living overseas in April 2000 and registered with a GP at some point in the April 2000 – April 2001 period have a slightly elevated level of fertility compared to the non-migrant (1971-1991) group. Overall, this model shows that the non-recent migrants who registered with a GP in the 1991-April 2000 period and were not living overseas 12 months before the 2001 census have a higher level of fertility compared to the non-migrants. Among the more recent migrant group (the final line in the model) there is a slightly higher level of fertility compared to the non-migrants. This model verifies the life table analysis which showed that the migrants from 1991-2001 had a higher level of fertility after the 2001 census.

Model 1B shows the benefit of this approach with the inclusion of an age covariate from the 2001 census form. The extent to which the different age profiles of the migrant and non-migrant groups change the likelihood of a birth in the period after the 2001 census are shown; the extent to which the higher level of migrant fertility because of the age of the migrant is clear. The model shows that for the different age groups there is a profile of coefficients which matches a typical age-specific fertility profile for England and Wales. For the coefficients among the different migrant groups the inclusion of age flattens the profile of these values towards the reference category (non-migrants) apart from for the migrants (1971-1991) where the odds ratio makes a modest increase from 1.1 to 1.2. For the migrant groups where the migrant was not overseas 12 months before the 2001 census there is a smaller change in the coefficient than for migrants who were overseas 12 months before where the coefficients decrease more drastically and decline below the reference category, non-migrants (1971-1991). This shows the relative importance of the age group of the migrant groups in accounting for the fertility trends among the migrant groups.

In Model 1C all the other covariates are included and a number of interaction terms. This model shows the fertility of migrants after the 2001 census controlling for socio-economic characteristics. For all the migrant groups, except the long-term migrants (1971-1991) and 2000-2001 migrants who were overseas 12 months before, there is a convergence towards the non-migrant group.

Table 8.17: Model 1: Discrete-time hazard model – migrants and non-migrants using the date of NHSCR registration and location of residence 12 months before the 2001 census, dependent variable – birth in 24 months after 2001 census

	Model A			Model B			Model C		
	B	Sig.	Exp(B)	B	Sig.	Exp(B)	B	Sig.	Exp(B)
Months since 2001 census									
1-6 months (Reference)	0.000	-----	1.000	0.000	-----	1.000	0.000		1.000
7-12 months	-0.085	0.002	0.919	-0.075	0.007	0.928	-0.045	0.101	0.956
13-18 months	-0.015	0.587	0.985	0.005	0.843	1.005	0.064	0.020	1.066
19-24 months	-0.125	0.000	0.882	-0.095	0.001	0.910	-0.013	0.653	0.987
Migrant indicator									
Not a migrant (1971-1991)	0.000		1.000	0.000		1.000	0.000		1.000
Migrant (NHSCR entry 1971-1991)	0.061	0.196	1.063	0.163	0.001	1.177	0.121	0.071	1.129
Migrant NHSCR entry 1991-April 2000, not overseas 12 mths before 2001	0.578	0.000	1.782	0.315	0.000	1.370	0.057	0.371	1.059
Migrant NHSCR entry 1991-April 2000, overseas 12 mths before 2001	0.163	0.646	1.177	-0.230	0.516	0.794	0.036	0.920	1.037
Migrant NHSCR entry April 2000-April 2001, not overseas 12 mths before 2001	0.932	0.000	2.538	0.637	0.000	1.890	0.412	0.000	1.510
Migrant NHSCR entry April 2000-April 2001, overseas 12 mths before 2001	0.201	0.127	1.223	-0.055	0.680	0.947	-0.090	0.534	0.914
Age group									
25-29 years (Reference)				0.000		1.000	0.000	.	1.000
15-19 years				-1.131	0.000	0.323	-0.691	0.000	0.501
20-24 years				-0.348	0.000	0.706	-0.037	0.540	0.963
30-34 years				-0.185	0.000	0.831	-0.231	0.000	0.794
35-39 years				-1.133	0.000	0.322	-1.268	0.000	0.282
40-44 years				-3.011	0.000	0.049	-3.392	0.000	0.034
Parity									
No children (Reference)							0.000		1.000
One child							0.398	0.000	1.489
Two or more children							-0.611	0.000	0.543
Age - Parity interaction									
15-19 x One child							0.046	0.673	1.047
20-24 x One child							0.158	0.022	1.171
30-34 x One child							-0.122	0.045	0.885
35-39 x One child							-0.112	0.177	0.894
40-44 x One child							-0.265	0.178	0.767
15-19 x Two or more children							1.084	0.000	2.958
20-24 x Two or more children							0.514	0.000	1.671
30-34 x Two or more children							-0.511	0.000	0.600
35-39 x Two or more children							-0.420	0.000	0.657
40-44 x Two or more children							0.326	0.093	1.386
Education level									
Level 2: 5+O levels, 5+CSEs (grade1), 5+GCSEs (grades A-C) (Reference)							0.000		1.000
Not recorded / other / missing / no value / other qualification							-0.605	0.000	0.546
No academic or professional qualifications							-0.126	0.057	0.882
Level 1: CSEs (grades 2-5), GCSEs (grades D-G)							-0.086	0.101	0.918
Level 3: 2+ A levels, 4 + AS levels							-0.065	0.339	0.937
Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND							-0.136	0.007	0.873

Age - Education interaction									
15-19 x No academic or professional qualifications							1.230	0.000	3.423
20-24 x No academic or professional qualifications							0.356	0.000	1.428
30-34 x No academic or professional qualifications							-0.051	0.587	0.950
35-39 x No academic or professional qualifications							-0.407	0.001	0.665
40-44 x No academic or professional qualifications							-0.891	0.002	0.410
15-19 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)							1.051	0.000	2.859
20-24 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)							0.215	0.010	1.239
30-34 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)							-0.105	0.155	0.901
35-39 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)							-0.248	0.014	0.781
40-44 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)							-0.203	0.384	0.816
15-19 x Level 3: 2+ A levels, 4 + AS levels							-0.267	0.111	0.765
20-24 x Level 3: 2+ A levels, 4 + AS levels							-0.241	0.014	0.786
30-34 x Level 3: 2+ A levels, 4 + AS levels							0.155	0.122	1.168
35-39 x Level 3: 2+ A levels, 4 + AS levels							0.244	0.062	1.276
40-44 x Level 3: 2+ A levels, 4 + AS levels							-0.099	0.772	0.906
20-24 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND							-0.716	0.000	0.489
30-34 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND							0.374	0.000	1.453
35-39 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND							0.728	0.000	2.070
40-44 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND							0.690	0.000	1.993
Economic position									
Employed Full Time (Reference)							0.000	-----	1.000
Not recorded / other / no value							0.322	0.000	1.379
Employed Part time							0.150	0.000	1.162
Self employed							0.186	0.001	1.205
Seeking work / Retired / Permanently Sick							0.087	0.066	1.091
Student							-1.000	0.000	0.368
Looking after home							0.529	0.000	1.697
Martial status									
Single (never married) (Reference)							0.000		1.000
Married (first marriage)							1.062	0.000	2.893
Re-married							1.060	0.000	2.885
Separated, Divorced, Widowed and other							0.511	0.000	1.666
Country of birth									
UK (Reference)							0.000		1.000
Ireland							0.124	0.378	1.132
India							-0.214	0.038	0.808
Pakistan							0.157	0.071	1.170
Bangladesh							0.148	0.123	1.160
East Africa							-0.091	0.524	0.913
Uncoded / other							-0.183	0.003	0.833
Constant	-5.452	0.000	0.004	-4.802	0.000	0.008	-5.323	0.000	0.005

Own elaboration based on ONS LS, March 2012.

8.8.2 Results from Model 2 – are similar factors influencing fertility among migrant and non-migrant groups?

Table 8.18 presents the results for Model 2 which examines whether the factors influencing fertility are similar among migrant and non-migrants groups. There are two migrant groups used in this analysis – those migrants who entered the LS with a registration with an NHS GP in the 1991-2001 period and those LS migrants who entered in this way plus those who arrived at the 2001 census for the first time (from migration date 2). A measure of the time between migration and the 2001 census is not included. The model shows that across each of the grouped 24 months in the May 2001-April 2003 period there is no period trend in the probability of giving birth in either the migrant or non-migrant groups.

The effect of age on the probability of giving birth is slightly different for migrants than for those who are continuously resident. Among non-migrants fertility is highest among those aged 20-29. In contrast to this, the migrant group has the highest odds of giving birth is the 20-24 years age group. Results by parity are similar for the non-migrants and the two migrant groups. Those women with a child have the highest likelihood to go on and have another birth in the 24 months after the 2001 census. This fits with the tight spacing of fertility which is typically observed in England and Wales (Ní Bhrolcháin, 1988).

However, there are noticeable differences when age and parity interaction terms are included. Among non-migrant teenagers there is a particularly high propensity to go on and have a higher parity birth (i.e. odds ratio of 2.8 for women with two or more children). This pattern will be discussed further in later models presented in this chapter. Among the migrant groups the age and parity interaction terms are not statistically significant and are generally similar across the two migrant groups except for the 40-44 age group – two child interaction. However, the non-migrant groups differ in direction with the 15-24 years age groups with parities 1 and 2 or more showing higher odds and the 35-44 years age group showing lower odds of a birth.

As a main effect in the analysis, education level is not a good predictor of fertility, possibly with the exception of where the education value was not recorded or uncodeable. However, the age and education interaction term shows there are large educational differentials in fertility that change over age. For the non-migrant group, the higher level of fertility among teenagers with no qualifications is statistically significant and shows that there is an important link between no qualifications and giving birth after the 2001 census (odds ratio of 3.4). This is also true for the 20-24 years age group, although the odds ratio for this age groups is substantially lower than for the 15-19 years age group. Among teenagers with low qualifications (level 1) there are also high levels of fertility in the subsequent 24 months. The delaying of fertility to later ages is evident among the non-migrant group with high levels of fertility among women in the 30-34 years and 35-39 years age groups in the level 3 and 4/5 education groups. Among migrant groups the results differ slightly. Teenage migrants with no qualifications are slightly less likely to have a birth, but among teenagers with a level 1 qualification there are comparable coefficients to the non-migrants. In the other age groups where the migrant has a low qualification (level 1) the coefficients are generally higher than for the corresponding non-migrants age groups. This is particularly the case for the 20-24 age group. In the higher qualification categories the coefficients for the migrants and non-migrants seem to differ (although to test formally the data would need to be pooled and migrant status / education interaction terms included). The younger age groups among the migrants have much higher coefficients than for the non-migrants, these cross over at the older ages showing the preference among older non-migrants to wait until later in life to give birth or that the migrant groups have already had their children and left these in their country of origin.

Table 8.18: Model 2: Discrete-time hazard model – non-migrants and migrants, no duration from migration variable, dependent variable – birth in 24 months after 2001 census

	Continually resident consistent 1991-2001			Migrants 1991-2001, NHSCR registration			Migrants 1991-2001, NHSCR registration and arrivals at 2001 census		
	B	Sig.	Exp(B)	B	Sig.	Exp(B)	B	Sig.	Exp(B)
Months since 2001 census									
1-6 months (Reference)	0.000	)	1.000	0.000		1.000	0.000		1.000
7-12 months	-0.061	0.046	0.941	0.108	0.239	1.114	0.066	0.439	1.069
13-18 months	0.070	0.020	1.072	0.056	0.552	1.058	0.061	0.492	1.063
19-24 months	-0.004	0.903	0.996	-0.049	0.626	0.952	-0.063	0.499	0.939
Age group									
25-29 years (Reference)	0.000	)	1.000	0.000		1.000	0.000		1.000
15-19 years	-0.675	0.000	0.509	-0.219	0.509	0.803	-0.270	0.382	0.763
20-24 years	-0.050	0.441	0.951	0.133	0.585	1.142	0.145	0.528	1.156
30-34 years	-0.248	0.000	0.780	-0.113	0.645	0.893	-0.238	0.309	0.788
35-39 years	-1.360	0.000	0.257	-0.743	0.029	0.476	-0.710	0.025	0.491
40-44 years	-3.433	0.000	0.032	-4.836	0.000	0.008	-2.966	0.001	0.052
Parity									
No children (Reference)	0.000	)	1.000	0.000		1.000	0.000		1.000
One child	0.391	0.000	1.479	0.390	0.006	1.476	0.398	0.003	1.488
Two or more children	-0.628	0.000	0.534	-0.723	0.000	0.485	-0.691	0.000	0.501
Age - Parity interaction									
15-19 x One child	0.094	0.420	1.098	-0.565	0.247	0.569	-0.605	0.209	0.546
20-24 x One child	0.202	0.007	1.224	-0.270	0.199	0.763	-0.286	0.160	0.751
30-34 x One child	-0.118	0.080	0.889	-0.113	0.582	0.893	-0.101	0.595	0.904
35-39 x One child	-0.115	0.210	0.891	0.133	0.669	1.143	0.043	0.883	1.044
40-44 x One child	-0.240	0.254	0.787	0.461	0.708	1.586	-0.681	0.461	0.506
15-19 x Two or more children	1.043	0.000	2.839	0.492	0.640	1.635	0.427	0.684	1.533
20-24 x Two or more children	0.571	0.000	1.769	0.017	0.955	1.017	0.014	0.962	1.014
30-34 x Two or more children	-0.591	0.000	0.554	0.118	0.637	1.125	0.086	0.709	1.090
35-39 x Two or more children	-0.455	0.000	0.634	0.195	0.564	1.216	0.124	0.691	1.132
40-44 x Two or more children	0.392	0.058	1.480	1.308	0.264	3.700	0.579	0.460	1.785
Education level									
Level 2: 5+O levels, 5+CSEs (grade1), 5+GCSEs (grades A-C) (Reference)	0.000	)	1.000	0.000		1.000	0.000		1.000
Not recorded / other / missing / no value / other qualification	-0.668	0.000	0.513	-0.439	0.018	0.645	-0.408	0.023	0.665
No academic or professional qualifications	-0.077	0.320	0.926	-0.422	0.034	0.656	-0.291	0.119	0.748
Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	-0.088	0.112	0.916	-0.435	0.122	0.647	-0.327	0.204	0.721
Level 3: 2+ A levels, 4 + AS levels	-0.026	0.715	0.974	-0.346	0.161	0.708	-0.372	0.120	0.690
Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	-0.125	0.024	0.882	-0.308	0.080	0.735	-0.295	0.076	0.745

Age - Education interaction									
15-19 x No academic or professional qualifications	1.215	0.000	3.370	0.990	0.018	2.691	1.028	0.009	2.796
20-24 x No academic or professional qualifications	0.394	0.000	1.484	0.374	0.171	1.454	0.329	0.204	1.390
30-34 x No academic or professional qualifications	-0.174	0.121	0.840	0.135	0.642	1.145	0.253	0.348	1.288
35-39 x No academic or professional qualifications	-0.480	0.001	0.619	-0.418	0.322	0.659	-0.288	0.438	0.749
40-44 x No academic or professional qualifications	-0.966	0.002	0.381	-	-	-	-0.693	0.575	0.500
15-19 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	1.046	0.000	2.847	1.057	0.046	2.877	1.024	0.038	2.785
20-24 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	0.201	0.023	1.223	0.642	0.106	1.900	0.430	0.253	1.537
30-34 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	-0.097	0.217	0.908	-0.348	0.443	0.706	-0.445	0.284	0.641
35-39 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	-0.188	0.079	0.828	0.184	0.732	1.201	-0.115	0.816	0.891
40-44 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	-0.182	0.447	0.834	-	-	-	-	-	-
15-19 x Level 3: 2+ A levels, 4 + AS levels	-0.289	0.101	0.749	0.107	0.876	1.113	0.043	0.949	1.044
20-24 x Level 3: 2+ A levels, 4 + AS levels	-0.283	0.007	0.754	-0.037	0.919	0.964	0.019	0.958	1.019
30-34 x Level 3: 2+ A levels, 4 + AS levels	0.088	0.426	1.091	0.156	0.671	1.169	0.403	0.242	1.496
35-39 x Level 3: 2+ A levels, 4 + AS levels	0.285	0.045	1.330	0.175	0.726	1.191	0.083	0.859	1.087
40-44 x Level 3: 2+ A levels, 4 + AS levels	-0.369	0.346	0.692	-	-	-	0.586	0.638	1.797
20-24 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	-0.880	0.000	0.415	0.125	0.682	1.133	0.021	0.944	1.021
30-34 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	0.343	0.000	1.410	0.216	0.396	1.241	0.431	0.072	1.539
35-39 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	0.751	0.000	2.120	0.253	0.458	1.287	0.253	0.423	1.288
40-44 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	0.661	0.002	1.936	2.076	0.060	7.975	0.667	0.435	1.948
Economic position									
Employed Full Time (Reference)	0.000		1.000	0.000		1.000	0.000		1.000
Not recorded / other / no value	0.292	0.000	1.339	0.468	0.000	1.598	0.438	0.000	1.550
Employed Part time	0.135	0.000	1.145	0.420	0.002	1.521	0.337	0.006	1.401
Self employed	0.187	0.002	1.205	0.326	0.129	1.386	0.255	0.187	1.291
Seeking work / Retired / Permanently Sick	0.046	0.390	1.047	0.164	0.310	1.179	0.128	0.390	1.137
Student	-1.140	0.000	0.320	-0.436	0.007	0.647	-0.560	0.000	0.571
Looking after home	0.530	0.000	1.699	0.479	0.000	1.615	0.455	0.000	1.576
Martial status									
Single (never married) (Reference)	0.000		1.000	0.000		1.000	0.000		1.000
Married (first marriage)	1.066	0.000	2.905	1.652	0.000	5.215	1.543	0.000	4.676
Re-married	1.085	0.000	2.959	1.339	0.000	3.817	1.223	0.000	3.396
Separated, Divorced, Widowed and other	0.502	0.000	1.652	1.090	0.000	2.975	1.048	0.000	2.852
Country of birth									
UK (Reference)	0.000		1.000	0.000		1.000	0.000		1.000
Ireland	0.410	0.020	1.507	0.251	0.337	1.286	0.098	0.681	1.103
India	-0.058	0.727	0.944	-0.250	0.124	0.779	-0.254	0.075	0.776
Pakistan	0.301	0.014	1.352	0.118	0.468	1.126	0.129	0.361	1.138
Bangladesh	0.299	0.014	1.348	0.137	0.456	1.147	0.086	0.600	1.089
East Africa	0.022	0.904	1.022	-0.193	0.469	0.825	-0.158	0.525	0.854
Uncoded / other	-0.241	0.001	0.786	-0.194	0.127	0.824	-0.146	0.147	0.864
Constant	-5.298	0.000	0.005	-5.722	0.000	0.003	-5.649	0.000	0.004

Own elaboration based on ONS LS, March 2012.

The impact of economic position on fertility differs for migrants and non-migrants: both migrant groups have higher odds ratio of fertility among women who are employed part time and also among students relative to the non-migrant group. In the last category 'looking after home' the higher coefficients could be related to women who have already given birth and short spacing of fertility. Fertility differentials according to marital status also differ between the migrant and non-migrant groups. While the married and remarried non-migrant groups have higher levels of fertility relative to the single group, (odds ratios of 2.91 and 2.96 respectively, for the migrant group where there was an NHSCR registration the corresponding coefficients are 5.2 and 3.8). This shows that marriage is more related to fertility among migrants than for non-migrants, as initially hypothesised in the methodology section. For the second migrant group, including arrivals at the 2001 census, the married odds ratios are 4.7 and for the remarried 3.4.

The final covariate is country of birth as recorded at the 2001 census. For the non-migrants, those women born in Ireland, Pakistan and Bangladesh have notably higher fertility than the UK born. Given this model controls for age and such a range of other socio-economic covariates this shows that long term, women from these countries of birth have a different propensity to give birth relative to women from the UK. Interestingly, among the migrant groups the coefficients for all three of these countries (Ireland, Pakistan and Bangladesh) actually decline relative to the non-migrant group and drop further when arrivals at the 2001 census are included.

8.8.3 Results from Model 3 – distinguishing between migrants overseas 12 months before the 2001 census: is there a fertility timing effect among recent migrants to England and Wales?

The purpose of this is analysis is to address research question three which is concerned with identifying whether there is a timing effect among recent migrants to England and Wales and to do so without using the NHSCR date of registration with a GP. Table 8.19 presents results for Model 3 which is an evolution of Model 1, but just considers the two migrant groups previously discussed and uses the 2001 census variable on whether the LS member was living overseas 12 months before. The headline finding from the inclusion of this variable is that the migrants to England and Wales in the 1991-2001 period who said that they were overseas 12 months before the 2001 census have a lower fertility in the May 2001-April 2003 period. Values for the migrant group which includes the arrivals at the 2001 census is slightly lower still (an odds ratio of 0.73 compared with 0.77 for the NHSCR migrants). This would be expected given earlier results on registration via the NHSCR and subsequent fertility. This suggests that there is actually a lower level of fertility in the 24 months after the migration event. The other covariates in the model show minor changes compared to Model 1. As identified for Model 2, women born in Ireland, Pakistan and Bangladesh have a higher level of fertility relative to UK born women.

Table 8.19: Model 3: Discrete-time hazard model – migrants, using flag for whether overseas 12 months before the 2001 census, dependent variable – birth in 24 months after 2001 census

	Migrants 1991-2001, NHSCR registration			Migrants 1991-2001, NHSCR registration and arrivals at 2001 census		
	B	Sig.	Exp(B)	B	Sig.	Exp(B)
Months since 2001 census						
1-6 months (Reference)	0.000		1.000	0.000		1.000
7-12 months	0.108	0.239	1.114	0.067	0.439	1.069
13-18 months	0.056	0.553	1.058	0.061	0.492	1.063
19-24 months	-0.048	0.629	0.953	-0.063	0.503	0.939
Migration status at 2001 census						
Not overseas 12 months before 2001 census (Reference)	0.000		1.000	0.000		1.000
Overseas 12 months before 2001 census	-0.258	0.051	0.772	-0.315	0.014	0.729
Age group						
25-29 years (Reference)	0.000		1.000	0.000		1.000
15-19 years	-0.208	0.531	0.812	-0.258	0.405	0.773
20-24 years	0.150	0.539	1.162	0.166	0.472	1.180
30-34 years	-0.110	0.656	0.896	-0.233	0.319	0.792
35-39 years	-0.747	0.028	0.474	-0.715	0.024	0.489
40-44 years	-4.816	0.000	0.008	-2.962	0.001	0.052
Parity						
No children (Reference)	0.000		1.000	0.000		1.000
One child	0.374	0.008	1.454	0.378	0.005	1.460
Two or more children	-0.740	0.000	0.477	-0.711	0.000	0.491
Age - Parity interaction						
15-19 x One child	-0.619	0.206	0.539	-0.665	0.168	0.514
20-24 x One child	-0.277	0.187	0.758	-0.297	0.143	0.743
30-34 x One child	-0.110	0.591	0.896	-0.098	0.605	0.906
35-39 x One child	0.152	0.627	1.164	0.061	0.832	1.063
40-44 x One child	0.443	0.719	1.557	-0.687	0.457	0.503
15-19 x Two or more children	0.435	0.680	1.544	0.361	0.731	1.434
20-24 x Two or more children	0.001	0.998	1.001	-0.003	0.991	0.997
30-34 x Two or more children	0.118	0.638	1.125	0.086	0.711	1.089
35-39 x Two or more children	0.210	0.536	1.233	0.139	0.656	1.149
40-44 x Two or more children	1.284	0.273	3.613	0.568	0.469	1.765
Education level						
Level 2: 5+O levels, 5+CSEs (grade1), 5+GCSEs (grades A-C) (Reference)	0.000		1.000	0.000		1.000
Not recorded / other / missing / no value / other qualification	-0.439	0.018	0.645	-0.407	0.023	0.666
No academic or professional qualifications	-0.427	0.032	0.652	-0.296	0.113	0.744
Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	-0.430	0.126	0.650	-0.323	0.209	0.724
Level 3: 2+ A levels, 4 + AS levels	-0.342	0.166	0.710	-0.369	0.122	0.691
Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	-0.289	0.100	0.749	-0.273	0.101	0.761

Age - Education interaction						
15-19 x No academic or professional qualifications	1.044	0.013	2.841	1.085	0.006	2.961
20-24 x No academic or professional qualifications	0.377	0.168	1.458	0.332	0.200	1.394
30-34 x No academic or professional qualifications	0.135	0.642	1.145	0.252	0.351	1.286
35-39 x No academic or professional qualifications	-0.423	0.316	0.655	-0.294	0.429	0.745
40-44 x No academic or professional qualifications	-	-	-	-0.687	0.579	0.503
15-19 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	1.038	0.050	2.825	1.004	0.042	2.728
20-24 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	0.650	0.101	1.916	0.437	0.245	1.548
30-34 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	-0.355	0.433	0.701	-0.455	0.275	0.635
35-39 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	0.176	0.742	1.193	-0.122	0.806	0.886
15-19 x Level 3: 2+ A levels, 4 + AS levels	0.152	0.825	1.164	0.085	0.900	1.089
20-24 x Level 3: 2+ A levels, 4 + AS levels	-0.038	0.918	0.963	0.021	0.952	1.021
30-34 x Level 3: 2+ A levels, 4 + AS levels	0.164	0.655	1.178	0.413	0.230	1.511
35-39 x Level 3: 2+ A levels, 4 + AS levels	0.170	0.733	1.186	0.084	0.858	1.088
40-44 x Level 3: 2+ A levels, 4 + AS levels	-	-	-	0.591	0.636	1.805
20-24 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	0.126	0.679	1.134	0.022	0.939	1.023
30-34 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	0.204	0.424	1.226	0.418	0.082	1.518
35-39 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	0.233	0.494	1.262	0.233	0.462	1.263
40-44 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	2.067	0.061	7.898	0.661	0.439	1.936
Economic position						
Employed Full Time (Reference)	0.000		1.000	0.000		1.000
Not recorded / other / no value	0.483	0.000	1.622	0.456	0.000	1.577
Employed Part time	0.428	0.001	1.535	0.347	0.005	1.415
Self employed	0.321	0.135	1.379	0.249	0.197	1.283
Seeking work / Retired / Permanently Sick	0.171	0.290	1.187	0.141	0.346	1.151
Student	-0.420	0.009	0.657	-0.544	0.000	0.580
Looking after home	0.498	0.000	1.645	0.478	0.000	1.613
Martial status						
Single (never married) (Reference)	0.000		1.000	0.000		1.000
Married (first marriage)	1.654	0.000	5.229	1.547	0.000	4.697
Re-married	1.343	0.000	3.832	1.226	0.000	3.409
Separated, Divorced, Widowed and other	1.085	0.000	2.961	1.044	0.000	2.841
Country of birth						
UK (Reference)	0.000		1.000	0.000		1.000
Ireland	0.259	0.323	1.296	0.106	0.656	1.112
India	-0.233	0.152	0.792	-0.235	0.100	0.790
Pakistan	0.121	0.457	1.129	0.130	0.358	1.139
Bangladesh	0.141	0.443	1.152	0.089	0.585	1.093
East Africa	-0.187	0.482	0.830	-0.151	0.543	0.860
Uncoded / other	-0.178	0.162	0.837	-0.128	0.202	0.880
Constant	-5.726	0.000	0.003	-5.652	0.000	0.004

Own elaboration based on ONS LS, March 2012.

8.8.4 Results from Model 4 – migrants (1991-2001) estimating duration from migration to the 2001 census: is there an elevated level of fertility after migration to England and Wales using a continuous measure of duration of residence?

The purpose of using these measures of duration from migration to England and Wales is to contribute towards answering question three which is concerned with whether there is an elevated level of fertility in the time immediately after migration to England and Wales. Model 4 develops Models 1 and 3 further, and like Model 3 uses a measure of the duration of residence in England and Wales. As explained in section 8.5.5, a range of estimates of the date of migration to England and Wales have been used. Model 4 shows results for migration dates 1 (the date of NHSCR registration) and 3 (the latest possible date of migration or minimum exposure to risk of birth). Table 8.20 shows the values for these migration dates. (Only migration dates 1 and 3 have been selected for inclusion because in the case of migration date 2 and 4 the distribution of the duration variable has been heavily modified by the assumptions used).

Looking first at migration date 1, the values for duration from migration are in the main not significant and there is very little difference between the different durations. It is possible that if there is a trend then it is a bimodal one with fertility concentrated among those women arriving between 0-3 years and 6+ years before April 2001. In the case of migration date 3 the results are similar to those for the NHSCR date of registration (migration date 1), reflecting that migration date 3 in the main uses the NHSCR registration date apart from in the years just before the 2001 census where the 12 month migration indicator was also used. Because of the inability of this variable and analyses so far to provide a firm indication of any duration effect for migrants, the next model will use an alternative, courser measure of date of migration and duration of residence, similar to that presented in Model 1.

Table 8.20: Model 4: Discrete-time hazard model – migrants, using duration from date of migration, dependent variable - birth in 24 months after 2001 census

	MIGRATION DATE 1 Migrants 1991-2001, NHSCR registration			MIGRATION DATE 3 Migrants 1991-2001, NHSCR registration and arrivals at 2001 census		
	B	Sig.	Exp(B)	B	Sig.	Exp(B)
Months since 2001 census						
1-6 months (Reference)	0.000		1.000	0.000		1.000
7-12 months	0.109	0.231	1.116	0.068	0.429	1.070
13-18 months	0.059	0.533	1.061	0.063	0.475	1.065
19-24 months	-0.046	0.648	0.955	-0.061	0.517	0.941
Duration from migration to 2001 census (April 2001)						
0-1 year since date of migration (Reference)	0.000		1.000	0.000		1.000
1-2 years since date of migration	-0.145	0.242	0.865	-0.160	0.164	0.852
2-3 years since date of migration	-0.059	0.628	0.942	-0.043	0.703	0.958
3-4 years since date of migration	-0.159	0.228	0.853	-0.154	0.208	0.857
4-5 years since date of migration	-0.119	0.385	0.888	-0.115	0.363	0.891
5-6 years since date of migration	-0.149	0.334	0.861	-0.109	0.445	0.897
6-7 years since date of migration	0.079	0.599	1.082	0.086	0.527	1.090
7-8 years since date of migration	-0.249	0.144	0.780	-0.231	0.152	0.794
8-9 years since date of migration	0.083	0.612	1.087	0.070	0.653	1.073
9-10 years since date of migration	-0.160	0.375	0.852	-0.133	0.432	0.875
Age group						
25-29 years (Reference)	0.000		1.000	0.000		1.000
15-19 years	-0.255	0.443	0.775	-0.302	0.329	0.739
20-24 years	0.121	0.618	1.129	0.133	0.565	1.142
30-34 years	-0.126	0.610	0.882	-0.256	0.273	0.774
35-39 years	-0.772	0.024	0.462	-0.737	0.020	0.479
40-44 years	-4.859	0.000	0.008	-2.997	0.001	0.050
Parity						
No children (Reference)	0.000		1.000	0.000		1.000
One child	0.396	0.005	1.485	0.402	0.003	1.494
Two or more children	-0.736	0.000	0.479	-0.706	0.000	0.493
Age - Parity interaction						
15-19 x One child	-0.493	0.314	0.611	-0.537	0.266	0.585
20-24 x One child	-0.254	0.228	0.775	-0.271	0.184	0.763
30-34 x One child	-0.112	0.587	0.894	-0.104	0.586	0.902
35-39 x One child	0.137	0.660	1.147	0.045	0.876	1.046
40-44 x One child	0.460	0.709	1.584	-0.691	0.456	0.501
15-19 x Two or more children	0.520	0.623	1.681	0.442	0.675	1.555
20-24 x Two or more children	0.063	0.837	1.065	0.057	0.844	1.059
30-34 x Two or more children	0.134	0.594	1.143	0.101	0.664	1.106
35-39 x Two or more children	0.211	0.533	1.236	0.133	0.67	1.142
40-44 x Two or more children	1.335	0.254	3.800	0.595	0.448	1.813
Education level						
Level 2: 5+O levels, 5+CSEs (grade1), 5+GCSEs (grades A-C) (Reference)	0.000		1.000	0.000		1.000
Not recorded / other / missing / no value / other qualification	-0.443	0.017	0.642	-0.406	0.023	0.666
No academic or professional qualifications	-0.429	0.032	0.651	-0.300	0.109	0.741
Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	-0.455	0.107	0.634	-0.340	0.187	0.712
Level 3: 2+ A levels, 4 + AS levels	-0.344	0.164	0.709	-0.375	0.117	0.687
Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	-0.322	0.067	0.725	-0.304	0.067	0.738

Age - Education interaction						
15-19 x No academic or professional qualifications	0.977	0.020	2.658	1.029	0.009	2.798
20-24 x No academic or professional qualifications	0.370	0.179	1.447	0.339	0.193	1.404
30-34 x No academic or professional qualifications	0.133	0.647	1.143	0.255	0.345	1.291
35-39 x No academic or professional qualifications	-0.397	0.347	0.672	-0.268	0.472	0.765
40-44 x No academic or professional qualifications	-	-	-	-0.668	0.589	0.513
15-19 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	1.079	0.042	2.941	1.046	0.035	2.847
20-24 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	0.663	0.095	1.941	0.447	0.235	1.563
30-34 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	-0.351	0.438	0.704	-0.447	0.283	0.639
35-39 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	0.204	0.704	1.226	-0.105	0.832	0.901
15-19 x Level 3: 2+ A levels, 4 + AS levels	0.096	0.889	1.101	0.043	0.95	1.044
20-24 x Level 3: 2+ A levels, 4 + AS levels	-0.056	0.878	0.945	0.011	0.975	1.011
30-34 x Level 3: 2+ A levels, 4 + AS levels	0.145	0.693	1.156	0.401	0.244	1.493
35-39 x Level 3: 2+ A levels, 4 + AS levels	0.168	0.738	1.182	0.083	0.861	1.086
40-44 x Level 3: 2+ A levels, 4 + AS levels	-	-	-	0.599	0.631	1.82
20-24 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	0.116	0.703	1.123	0.014	0.963	1.014
30-34 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	0.223	0.382	1.249	0.443	0.065	1.558
35-39 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	0.285	0.402	1.330	0.283	0.371	1.327
40-44 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	2.081	0.059	8.012	0.679	0.427	1.971
Economic position						
Employed Full Time (Reference)	0.000		1.000	0.000		1.000
Not recorded / other / no value	0.465	0.000	1.591	0.439	0.000	1.551
Employed Part time	0.422	0.001	1.524	0.339	0.006	1.403
Self employed	0.329	0.126	1.390	0.254	0.189	1.290
Seeking work / Retired / Permanently Sick	0.160	0.324	1.173	0.123	0.408	1.131
Student	-0.435	0.007	0.647	-0.557	0.000	0.573
Looking after home	0.483	0.000	1.620	0.463	0.000	1.588
Martial status						
Single (never married) (Reference)	0.000		1.000	0.000		1.000
Married (first marriage)	1.653	0.000	5.224	1.548	0.000	4.702
Re-married	1.339	0.000	3.815	1.226	0.000	3.408
Separated, Divorced, Widowed and other	1.099	0.000	3.000	1.054	0.000	2.870
Country of birth						
UK (Reference)	0.000		1.000	0.000		1.000
Ireland	0.245	0.351	1.277	0.114	0.634	1.121
India	-0.260	0.112	0.771	-0.229	0.114	0.796
Pakistan	0.100	0.543	1.105	0.137	0.340	1.146
Bangladesh	0.129	0.486	1.138	0.102	0.539	1.107
East Africa	-0.190	0.475	0.827	-0.125	0.616	0.882
Uncoded / other	-0.199	0.121	0.820	-0.118	0.251	0.889
Constant	-5.625	0.000	0.004	-5.600	0.000	0.004

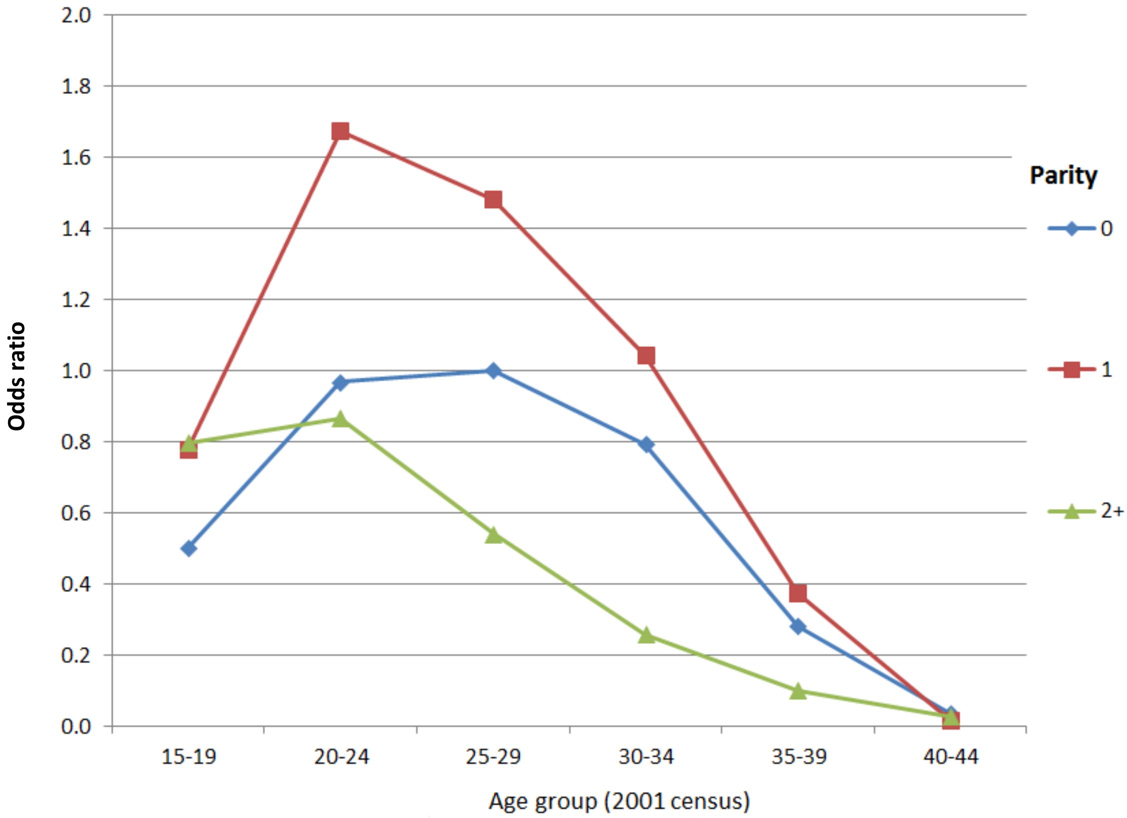
Own elaboration based on ONS LS, March 2012.

8.8.5 Results from Model 5 – estimating migrant fertility for migrant groups using the date of NHSCR registration and 12 month migration indicator at the 2001 census, what about non-recent migrants?

Model 5 reduces the migrant categories in use and is concerned with using the 12 month migration indicator from the 2001 census to identify migrants to England and Wales in comparison with the longer-term migrants (migrant (1971-1991)) and non-migrants. The purpose of this analysis is to identify if there is a different risk of birth for recent migrants to England and Wales using broader measures of duration from migration and to compare the odds of a birth with non-recent migrants and non-migrants since the start of the LS. Again, this is seeking to identify if there is any discernible migration-birth timing effect for births after the 2001 census. This differs from the past models because of the way in which both non-migrants and non-recent migrants (1971-1991) are included in the model as comparison groups. Again, the recent migrant group in this case (migrant April 2000-April 2001 (overseas 12mths before 2001 census)) shows a lower likelihood of giving birth relative to the non-migrants (1971-1991) and the other migrant groups (odds ratio of 0.92). There is hardly any difference in coefficients between the longer term migrants (1971-1991) and the 1991-2000 migrants (both have odds ratios of 1.18).

One aspect of the analysis in this chapter which was touched upon earlier is the higher level of fertility at certain ages and parities. In particular, in section 8.8.2 the high fertility of teenagers at higher parities was discussed. It is possible to extend the analysis of this further by using the interaction terms from this final model to show the estimated odds ratios of birth in the 24 months after the 2001 census. The model used the 25-29 years age group and parity 0 as the reference categories. Figure 8.4 shows the relative risk of a birth for each of the age groups and parities as of the 2001 census. In the teenage group the highest risk of a birth is for those teenagers who have already had one or two children prior to the 2001 census. The likelihood of a birth at parity zero loosely follows the age-specific fertility rate profile for England and Wales with a peak in the 25-29 years age group after which there is a decline to the older age groups. Relative risk of a birth at parity one is highest among the 20-24 years age group with a steep decline thereafter. For parity two the profile is high for the teenage age group, rising for the 20-24 years age group, thereafter declining. Among datasets covering England and Wales the ONS LS is probably one of the few able to disaggregate such clear trends by age and parity.

Figure 8.4: Estimated odds ratios of a birth in 24 months following the 2001 census for age and parity interaction terms, Model 5



Own elaboration based on ONS LS, March 2012.

Table 8.21: Model 5: Discrete-time hazard model – migrants and non-migrants using the date of NHSCR registration and location of residence 12 months before the 2001 census, dependent – variable birth in 24 months after 2001 census

	B	Sig.	Exp(B)
Months since 2001 census			
1-6 months (Reference)	0.000		1.000
7-12 months	-0.046	0.099	0.955
13-18 months	0.064	0.020	1.066
19-24 months	-0.013	0.647	0.987
Migrant indicator			
Not a migrant (1971-1991)	0.000		1.000
Migrant (1971-1991)	0.166	0.013	1.181
Migrant 1991 - April 2000 (not overseas 12mths before 2001 census)	0.167	0.004	1.182
Migrant April 2000 - April 2001 (overseas 12mths before 2001 census)	-0.081	0.541	0.922
Age group			
25-29 years (Reference)	0.000		1.000
15-19 years	-0.692	0.000	0.501
20-24 years	-0.033	0.588	0.967
30-34 years	-0.233	0.000	0.792
35-39 years	-1.269	0.000	0.281
40-44 years	-3.393	0.000	0.034
Parity			
No children (Reference)	0.000		1.000
One child	0.393	0.000	1.482
Two or more children	-0.616	0.000	0.540
Age - Parity interaction			
15-19 x One child	0.043	0.695	1.044
20-24 x One child	0.154	0.026	1.166
30-34 x One child	-0.119	0.051	0.888
35-39 x One child	-0.107	0.196	0.898
40-44 x One child	-0.261	0.184	0.770
15-19 x Two or more children	1.077	0.000	2.937
20-24 x Two or more children	0.505	0.000	1.657
30-34 x Two or more children	-0.509	0.000	0.601
35-39 x Two or more children	-0.416	0.000	0.660
40-44 x Two or more children	0.331	0.089	1.393
Education level			
Level 2: 5+O levels, 5+CSEs (grade1), 5+GCSEs (grades A-C) (Reference)	0.000		1.000
Not recorded / other / missing / no value / other qualification	-0.605	0.000	0.546
No academic or professional qualifications	-0.134	0.043	0.875
Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	-0.085	0.103	0.918
Level 3: 2+ A levels, 4 + AS levels	-0.070	0.305	0.933
Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	-0.139	0.006	0.870

Age - Education interaction			
15-19 x No academic or professional qualifications	1.244	0.000	3.470
20-24 x No academic or professional qualifications	0.361	0.000	1.435
30-34 x No academic or professional qualifications	-0.045	0.637	0.956
35-39 x No academic or professional qualifications	-0.399	0.001	0.671
40-44 x No academic or professional qualifications	-0.882	0.002	0.414
15-19 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	1.055	0.000	2.873
20-24 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	0.211	0.011	1.236
30-34 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	-0.105	0.155	0.901
35-39 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	-0.248	0.013	0.780
40-44 x Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	-0.203	0.384	0.816
15-19 x Level 3: 2+ A levels, 4 + AS levels	-0.261	0.120	0.771
20-24 x Level 3: 2+ A levels, 4 + AS levels	-0.238	0.015	0.788
30-34 x Level 3: 2+ A levels, 4 + AS levels	0.159	0.114	1.172
35-39 x Level 3: 2+ A levels, 4 + AS levels	0.249	0.057	1.283
40-44 x Level 3: 2+ A levels, 4 + AS levels	-0.095	0.781	0.909
20-24 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	-0.715	0.000	0.489
30-34 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	0.375	0.000	1.455
35-39 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	0.730	0.000	2.076
40-44 x Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	0.693	0.000	1.999
Economic position			
Employed Full Time (Reference)	0.000		1.000
Not recorded / other / no value	0.328	0.000	1.388
Employed Part time	0.153	0.000	1.165
Self employed	0.186	0.001	1.205
Seeking work / Retired / Permanently Sick	0.086	0.068	1.090
Student	-1.003	0.000	0.367
Looking after home	0.530	0.000	1.700
Martial status			
Single (never married) (Reference)	0.000		1.000
Married (first marriage)	1.062	0.000	2.893
Re-married	1.060	0.000	2.886
Separated, Divorced, Widowed and other	0.510	0.000	1.666
Country of birth			
UK (Reference)	0.000		1.000
Ireland	0.067	0.634	1.070
India	-0.276	0.007	0.759
Pakistan	0.108	0.213	1.114
Bangladesh	0.097	0.314	1.102
East Africa	-0.139	0.331	0.871
Uncoded / other	-0.228	0.000	0.796
Constant	-5.323	0.000	0.005

Own elaboration based on ONS LS, March 2012.

8.8.6 Results from Model 6 – estimating fertility for only those migrants overseas 12 months before the 2001 census

Table 8.22 shows the final discrete-time hazards model which has been specified. This uses the sample of ONS LS members at the 2001 census who said that 12 months before the census they were living overseas. In total for the 24 month period of observation there are just 69 births to these LS members (and 14,596 person months of exposure). Therefore, because of observed zeros, variables for many categories specified have been collapsed. Inclusion of the same set of variables is again important in allowing comparability between the regression models. In this case the Rendall and Ball country of birth grouping has been included in the model because the use of a grouping using the other countries of birth specified in past models was not possible. Indeed, in this case, for the sub-sample of migrants selected the Rendall and Ball country grouping shows that migrants from low income countries who were living overseas 12 months before the 2001 census are 2.2 times more likely to give birth in the period of observation compared to migrants from high income countries. This result is statistically significant at the 95% level. In addition to this, there is a clear parity effect with those migrants who already had one child being 2.1 times more likely to give birth again than those persons who do not already have a child (only main effects have been included because of the reduced sample size). The strongest effect is for marital status with those persons who were married being 10 times more likely to give birth compared to the unmarried (although it has not been possible to disaggregate the un-married into different groups which may reduce the effect of the married).

Table 8.22: Model 6 – Discrete-time hazards model – sample of migrants who were overseas 12 months before the 2001 census, dependent variable – birth in 24 months after 2001 census

	SAMPLE OF MIGRANTS OVERSEAS 12 MONTHS BEFORE 2001 CENSUS		
	B	Sig.	Exp(B)
Months since 2001 census			
1-6 months (Reference)	0.000		1.000
7-12 months	0.032	0.915	1.032
13-18 months	-0.857	0.037	0.424
19-24 months	-0.137	0.681	0.872
Age group			
25-29 years (Reference)	0.000		1.000
15-19 years	-0.423	0.488	0.655
20-24 years	0.094	0.784	1.098
30-34 years	-0.710	0.058	0.492
35-39 years	-0.612	0.208	0.542
40-44 years	0.000	0.006	0.000
Parity			
No children (Reference)	0.000		1.000
One child	0.755	0.013	2.127
Two or more children	-0.544	0.265	0.581
Education level			
Level 2: 5+O levels, 5+CSEs (grade1), 5+GCSEs (grades A-C) (Reference)	0.000		1.000
Not recorded / other / missing / no value / other qualification	-0.572	0.430	0.565
No academic or professional qualifications	-0.221	0.659	0.802
Level 1: CSEs (grades 2-5), GCSEs (grades D-G)	-0.769	0.290	0.464
Level 3: 2+ A levels, 4 + AS levels	-0.139	0.792	0.870
Level 4/5: First degree, Higher degree, NVQ levels 4-5, HNC, HND	-0.518	0.275	0.596
Economic position			
Employed	0.000		1.000
Not employed / economically inactive	-0.112	0.701	0.894
Martial status			
Unmarried	0.000		1.000
Married	2.384	0.000	10.847
Rendall and Ball (2004) - Country of birth			
High income	0.000		1.000
Missing	0.205	0.572	1.227
Low income	0.797	0.042	2.219
Constant	-6.382	0.000	0.002

Own elaboration based on ONS LS, December 2012.

8.8.7 Summary

Section 8.7 discussed the life table results showing relatively higher risk of birth among migrants relative to non-migrants. The primary aim of this section of the chapter has been to estimate whether migrants to England and Wales have an elevated level of fertility in the period after their migration and to identify whether this relates to the selection into migrant status of women who are at a higher risk of childbearing, or whether there is elevated fertility associated with the timing of the migration event. A series of discrete-time hazards models have been presented, each taking an alternative approach to the identification of groups of recent migrants.

Model 1 was an evolution of the analysis which was shown in the life tables comparing migrant and non-migrant fertility. The models which composed this showed that migrants to England and Wales in the 1991-2001 period have a higher risk of birth in the 2001-2003 period without controlling for other covariates (Mode1 1A). When age was included in Model 1B this showed that the highest fertility is among the 25-29 years age group and that among the oldest and the youngest groups there are much lower coefficients. Migrants who were overseas 12 months before the 2001 census in this model had a lower likelihood of giving birth than non-migrants. The 'registration effect' of late NHSCR entry associated with conception / pregnancy was identified in this model. The higher level of fertility shown in the life table analysis and in Model 1A is related to the age structure of migrant groups (Model 1B) and their socio-economic characteristics (Model 1C).

In Model 2 the factors affecting fertility among migrant and non-migrant groups were compared. The constant terms in this model showed the differences in the risk of a birth in the 24 months after the 2001 census for the different groups. Comparison of the effects of covariates for migrant and non-migrant groups showed differences in terms of the effects of age and parity, and age and education interaction terms. With the age parity interaction term the degree to which there is a high level of fertility among teenage women who have already started their childbearing is clear.

Through use of the indicator on whether the LS member was overseas 12 months before the 2001 census, Model 3 sought to remove any NHSCR registration effect from the analysis and simplify the modelling of duration from migration to birth. Again, the primary aim of this analysis was to contribute towards answering research question 3; to identify any timing effect among recent migrants. The purpose of this model was to simplify the identification of the recent migrants who were overseas just before the 2001 census. For the migrant groups who we truly believe were overseas 12 months prior to the 2001 census, the risk of fertility was actually lower in the two years after the 2001 census relative to those LS members who were not overseas 12 months before. This could be related to the type of migrant to England and Wales recorded in the time period used in the analysis here – asylum applications increased in the late 1990s (Matz et al., 2001). The reasons for migrating and the experience of arrival in England and Wales may be related to the subsequent family building.

Model 4 which used the duration from the NHSCR date of registration and also the most recent date of migration could not identify a clear, or statistically significant, difference in the risk of a birth between recent and non-recent migrants. Following on from this, Model 5 used alternative groupings of migrants and non-migrant comparison groups to try and distinguish a trend. Model 5 could not identify an elevated level of fertility among migrants who reported that they were overseas 12 months before the 2001 census; indeed coefficients were similar to the non-migrant reference group. The group reporting that they were overseas 12 months before the 2001 census and registered on the NHSCR April 2000-April 2001 showed a very slightly lower likelihood of giving birth relative to the non-migrants.

Model 6 separately analysed fertility for migrants overseas 12 months before the 2001 census, finding that, in this case alone across the models presented, the link with being from a low income country is a valid predictor among this group.

8.9 Conclusions

This chapter has sought to estimate the fertility of ONS LS members who entered the dataset in the 1991-2001 period and were resident at the 2001 census compared with LS members who remained resident in England and Wales between census dates. The three research questions were, firstly whether recent migrants show an elevated risk of fertility compared to non-migrants; secondly, whether the age and or socio-economic composition of the migrant groups can explain the elevated fertility among recent migrants compared to non-migrants; and thirdly, whether, once socio-economic composition of migrants is controlled for, there is an elevated level of fertility associated with the migration event itself.

With regard to the first question, migrants do have a higher risk of a birth compared to non-migrants. In section 8.5 life table results showed the migrant group had a higher hazard of a birth in the 24 months following the census. The cumulative hazard through the time period, the survivor function, showed that migrants are twice as likely to have given birth during the interval compared to non-migrants. This was consistent through the period of observation, even as one moves away from the first 9 months during which time there could be the effect of the late registrations of migrants when pregnant. Model 1A in section 8.8 also identified this higher likelihood of a birth.

In relation to the second question, the subsequent discrete-time hazard models shows that the migrant group has a higher risk of birth because of the socio-economic characteristics of the group. In the first discrete-time hazards model specified, Model 1, the regression coefficients also showed this higher risk of a birth for migrants in the same time period, verifying the life table outputs. Subsequently, the inclusion of the other covariates decreased the strength of the migrant category on the likelihood of a birth. As identified in the analysis of the descriptives of the migrant and non-migrant groups in section 8.4, the differences in the socio-demographics of the migrant and non-migrant groups have important implications for the analysis. These are features which the life table analysis did not take into account. The outcomes of this were seen in the discrete-time hazards models which showed that when these socio-demographic characteristics are controlled for, the migrant group still has an elevated level of fertility, but not substantially higher compared to non-migrants in the timeframe of interest. Results for Model 1 showed that the characteristics of the migrant group are important in accounting for the higher fertility in the period after the 2001 census rather than any timing trend (included in later models).

Therefore, with reference to the question on the timing of fertility with reference to the migration event, this analysis has identified that among migrants to England and Wales in the period between 1991 and 2001 there was a higher risk of birth after the 2001 census, but this was largely because of the different socio-economic characteristics of the migrant group, particularly the age profile, marital status and country of birth of the migrants. No clear duration / timing effect similar to that identified in past research was discernible from this analysis. The use of the 2001 census variable on the 12 month migration indicator was also used in the analysis to stand in comparison alongside the NHSCR registration date. This did not find a link between being a 'genuinely recent' migrant and subsequent elevated fertility. In separate analysis of those migrants overseas 12 months before the 2001 census, those from low income countries were much more likely to give birth once other characteristics were controlled for.

One of the interesting by-products of this analysis has been the identification of the high coefficients for teenagers who are already at parities of one or more as of the 2001 census. The plots in Figure 8.4 showed that among parities of two and above there is a decline in the odds from the younger ages. Interaction terms from the models specified showed that there is a relatively higher risk of birth in the teenage years when there has already been one or more births. In addition to this, the age and education interaction terms showed that having minimal or no educational qualifications is associated with a much higher risk of birth. With regard to the migrant groups there was a different picture with women who are highly qualified and in slightly younger age groups being likely to give birth while older age groups are where the non-migrant highly qualified fertility is concentrated.

It is important to recall that in Chapter 4 as part of the regression analysis on socio-economic factors associated with attrition from the study, being non UK-born was associated with drop out from the LS. Because of this only the 24 months after the 2001 census have been used in this analysis. If the likelihood of making an unrecorded embarkation from the LS is indeed higher for foreign born LS members then it is possible that the life table and regression analysis in this chapter could be underestimating the true level of fertility among migrants. This would be the case because the number of women exposed to risk of birth in the samples selected could be higher than the number actually resident. Births would be coming from only those women remaining resident from the 2001 census. The degree of attrition in the LS and the possible impacts of this on the present analysis will not be fully known until after the linkage and processing of the 2011 census data.

At the 2011 census the month and year of first migration to England and Wales was asked on the census form for the first time since 1971. Given the higher levels of migration in the period since the mid-1990s and the way in which the ONS LS does not record the true date of migration this is a very important enhancement to the dataset. In addition to making analysis of the date of migration and subsequent fertility clearer for future analysis, within which there has been a consistent rise in the fertility rates in England and Wales, this question also raises interesting analysis possibilities related to the functioning of the ONS LS and the NHSCR. In the next chapter a full discussion of the analysis in this thesis is provided and conclusions are drawn.

Chapter 9

Conclusions

This chapter discusses the main contributions made in this thesis and contextualises this work within the existing literature on the fertility of migrants and demographic research using the Office for National Statistics (ONS) Longitudinal Study (LS). The first aim of this thesis was to identify if migrants to England and Wales exhibit an elevated level of fertility shortly after their migration and the reasons for any association. A second aim was to accurately account for non-continually resident LS members between census dates and use these LS members in fertility analysis. In many ways these aims are inter-related; the research in the latter part of this thesis on the fertility of migrants to England and Wales is only possible because of the detailed selection of LS members based on the information recorded on them in the LS, which was used in the typologies devised.

Within demographic research in England and Wales there has been little work on the timing of fertility among migrants. Because the ONS LS combines census, NHS and vital registration (births) data, it was identified as an appropriate dataset for the analysis of migrant fertility, provided the sample exposed to risk of birth could be established and appropriate measures of the duration from migration to England and Wales used. A key aim of this work was to account for the non-continually resident LS members between census dates and use these LS members in fertility analysis; this aspect came about because the changing sample of the ONS LS year on year is often not recognised in research using the data. Measures of the quality of the ONS LS relate to single years but do not take into account the residence of the LS member in the decade as a whole. Therefore, the first two chapters of analysis were concerned with the identification of samples from the ONS LS which reflect the changing composition of the population of England and Wales through time with the inclusion of LS members who were not continually resident, a group predominantly made up of migrants to and from England and Wales. Past research has not taken such a fine-grained approach to including LS members not continually resident in the dataset. Results in Chapter 5 showed that the inclusion of non-continually resident LS members, a group mainly composed of migrants, led to fertility rates from the ONS LS which were more closely matched to England and Wales figures. This highlights the importance of the year on year changes in the population of England and Wales recorded in the LS.

With the main group of non-continually resident LS members being migrants, the contribution from this group is important to include. Research like that of Portanti and Whitworth (2009) using LS members continually resident is selecting a sample which has a lower fertility rate anyway in the context of England and Wales and is not representative of the population through time. This is not uncommon though; generally research using the LS could acknowledge that the dataset is not completely 'longitudinal' and not a complete 'study'. To elaborate on the longitudinal point further, there is attrition in the LS which is often not recognised even though this is systematic and for several decades there is a cumulative influence in these trends for the profile of LS members selected (i.e. as shown in Chapter 4, certain socio-economic characteristics are strongly associated with attrition). To elaborate on the study point, this research identified that exposure to risk of birth is a function of fertility or planned fertility and the date of migration does not seem to be recorded by the NHSCR. Therefore, the ability to over-estimate the fertility of migrants or under-estimate their exposure to risk, as illustrated in Chapter 6, is a risk. Additionally, cancellations, where persons are removed from patient registers because they have not responded to notifications, are recorded by the NHSCR but not included in the LS (if this data were included in the LS there could be difficulties as there is no uniform process through which persons are removed from the local patient register).

For the analysis in chapters following Chapter 4, having a set of residence typologies allowed identification of migrants with far greater ease. This was shown in Chapter 6, which was concerned with the identification of samples of ONS LS members who migrated to England and Wales. Potential bias in the LS with the registration of migrants at the time when they wish to conceive or become pregnant was a cause for concern. This particular chapter was therefore concerned with identifying when migrants give birth relative to their date of entry onto the NHSCR, which in the LS is referred to as the date of migration to England and Wales. It was identified that registration with a General Practitioner (and entry to the NHSCR and the LS) is a function of fertility. Therefore, using this date as the date of migration to England and Wales would bias estimates of the duration from entry to the ONS LS to subsequent birth; leading to a potentially false picture of a short duration from 'migration' (but in this case entry to the ONS LS) to birth. This finding is one that should be considered for future LS research using the NHSCR date of registration.

As a result of this finding, the migration indicator from the 2001 census was used along with the date of registration on the NHSCR in Chapter 7 to identify samples of migrants entering the ONS LS between 1991 and 2001. Again, analysis in this chapter identified the late NHSCR registrations, those persons stating at the 2001 census that they were not overseas 12 months before but registered with a GP in the April 2000-April 2001 period, had an extremely high fertility rate in 2001.

Subsequently, Chapter 8 used two approaches to identifying migrants. The first was the use of information on the registration of the female LS member and the indicator on whether the LS member was overseas 12 months before the 2001 census to estimate the duration from migration to the 2001 census. The second approach was the use of a scheme of recent migrants, non-recent migrants and non-migrants. Three research questions were posed. Firstly, do migrants show a higher level of fertility in the period after the 2001 census compared to non-migrants? Secondly, does the socio-economic composition of the migrant group lead to elevated fertility compared with non-migrants? Thirdly, is there an elevated level of fertility associated with the migration event itself; some form of timing effect? Life table and regression results showed that migrants (1991-2001) have an elevated level of fertility when compared to non-migrants. Analysis using discrete-time hazards models, progressively adding socio-economic information showed that higher fertility among migrants arises because of their socio-economic profile. In summary, the analysis showed an elevated risk of a birth for migrants because of their socio-demographic profile. Therefore, at the current time it is not possible to prove that there is a timing effect to the period fertility rate caused by the recent migrants to England and Wales. This differs from the findings of Toulemon (2004) and Andersson (2004) where it was identified that recent migrants show a short duration from the date of migration to subsequent fertility. However, Toulemon (2004) identified post-arrival elevated fertility with age-specific fertility schedules alone. In contrast, Andersson (2004) used a range of demographic variables as covariates but without marital status or socio-economic characteristics. Like the work of Mussino and Van Raalte (2008) on fertility among migrants to Italy and Russia, the findings here do not support the presence of disruption to fertility or family formation being part of the post-migrant fertility patterns for female migrants to England and Wales.

A by-product of the analysis was the identification of the high fertility of teenagers in the ONS LS. Age and parity interaction terms showed that there is a high propensity of teenagers at the 2001 census who have already had one or more children to go on and have another in the 24 months after the 2001 census. While the coefficients may seem high, it is important to recall that this is a two year time frame and that the age at which childbearing begins is key to subsequent fertility patterns. Among datasets for England and Wales, the ONS LS is one of the few that could express such a trend among teenagers. As demonstrated, the exposure to risk of birth for migrants in the LS was closely associated with the propensity to give birth. This made disentangling migration and fertility complex; an issue that was overcome as much as possible in the analysis.

As stated in the introduction, the linkage of 2011 census data into the ONS LS is on-going. The range of variables from the census along with the closure of the 2001-2011 decade will unlock new research opportunities when the refreshed LS becomes available. The migrant fertility element of this thesis looked at the period directly after the 2001 census because of uncertainty about unobservable attrition in the years further away from the census and the need for vital covariate information (such as the country of birth of the migrant) from the 2001 census. For the estimation of the fertility of migrants in the period from 2001-2011, some of the work that has been completed in this thesis would not be necessary. This is because a question was included in the 2011 census asking persons born overseas for the month and year of arrival in the UK. Although this provides an important set of data on the true date of migration, it would be preferable if this had been specific to each country making up the UK, as it is not possible to distinguish between migrants who, for example, first arrived in Scotland and then moved to England and Wales at a later date. For example, research on the date of migration to England and Wales could have an NHSCR registration in September 2005, when the LS member arrived in Scotland in March 2004 for the first time. Therefore, at the 2011 census they would correctly record that they first arrived in the UK for the first time in March 2004. In reality, it is likely that only a minority of international migrants who are LS members move within the UK, but this is something to consider in such work and that the ONS LS is specific to England and Wales.

Including the question on the date of migration will allow the estimation of an exact duration from migration and the relationship between this and subsequent fertility should be clearer. In particular, having more precise information in the LS on the true date of migration will be extremely valuable in the context of a decade of high migration, as shown in Chapter 2. Among datasets covering England and Wales the LS will be one of a few with a sufficiently large number of cases of this type to have the potential to understand the fertility of these migrants to England and Wales. In the present analysis migrants from the ‘A8’ countries could not be studied because of the identification of the association between pregnancy and LS entry, unobservable attrition in the LS (in particular the way in which this is associated with being foreign born (see Chapter 4) and the lack of covariate information from a census. This was a major limitation with this work and the timeframe of research (2001-2003) in Chapter 8 represents the low point of recent fertility rates in England and Wales.

Therefore, the first possibility for future research would centre on use of the 2011 census question on the date of migration. Comparing responses to this against the date of GP registration recorded by the NHSCR and the question on the place of residence 12 months before the 2011 census would allow identification of the lag between migration and GP registration. Using a series of descriptive statistics and regression models the relative socio-economic / demographic trends for different countries of birth in registration with a GP could be estimated. This would provide a sound methodological / data quality output which would be important in the context of future demographic research using migrants from the fourth decade of the LS.

The second possibility is to use the findings from the use of the 2011 census question on the date of migration to calculate the exact exposure from the date of migration to subsequent birth. Therefore, a study of the timing of childbearing among 2001-2011 migrants could be made. This encompasses the post-2004 A8 migration to England and Wales. In addition to the migration information from the 2011 census, the 2011 census data would provide valuable information on pre- and post-migration childbearing as for non-LS members in the same household we have country of birth information. Through use of the own-children method timing of childbearing can be understood as a whole, accounting for pre- and post-migration fertility. In the context of the rise in fertility from 2001 the full contribution of the timing of fertility among migrants to England and Wales who were at the 2011 census should become clear. This would be a very important contribution to the material on migration and fertility interrelations.

Along with the availability of the refreshed ONS LS in 2013, there are changes being made to the provision of user support, with the Economic and Social Research Council (ESRC) funding a central 'hub' to make the longitudinal studies in the UK available from one central body. This provides a good opportunity for the collation of materials on the functioning, construction, past research and data quality in one place. At the current time this material is spread between the ONS and the Centre for Longitudinal Study Information and User Support (CeLSIUS). Some of the best diagrams on the operation of the LS and the combination of data only became clear from presentations delivered by ONS LS staff at the ONS. Collecting this in one place for those wishing to work with the ONS LS would be helpful in clarifying the functioning of the dataset for its users and encouraging new ideas on the use of the dataset. Documentation on the complex functioning of the ONS is also in need of updating. The most recent comprehensive volume on this is still Hattersley and Creeser (1995), since which there have been two censuses and changes to the construction of the dataset in the early 1990s and more recently, in 2008, the separation of the NHSCR from the ONS related to the Central Health Register Inquiry System (CHRIS) becoming the Patient Demographic System (PDS). With the linkage of the 2011 census data underway, documentation of the current processes used to combine data in the LS would be timely and could facilitate further research like this which wishes to make use of the administrative data which is combined in the LS.

More generally, within social sciences research at the current time and also connected to uncertainty over whether there may be a census in 2021, the LS may become an increasingly important source of data if the linkage of additional administrative data into the LS is facilitated. However, the degree to which the census dates are crucial in accounting for the members of the LS seems to be missed in some arguments in this area. Without the set datum which the census provides, data matching and linking has the potential to become meaningless without certainty over the relative reliability of each source and a fixed datum to connect to. Other data would need to be linked to bring reliability to the LS and the range of socio-economic information collected at the decennial census.

In closing, this analysis has identified that non-continuously resident LS members form an important part of the dataset which are referred to as part of the 1% sample which the LS constitutes, but is often not included in demographic analysis using the LS. Detailed research into the sample of migrants in the LS and their exposure to risk identified that exposure to risk is a function of fertility meaning that the date of NHSCR registration is not reliable as a proxy for the date of migration. The final analysis in this thesis suggests that the way in which migrants have consisted of women in the key childbearing ages and their other socio-economic characteristics which has led to the higher fertility of migrants rather than a strong fertility timing preference among recent migrants to England and Wales. Results from this analysis do not suggest that there is disruption to fertility from migration, or that there is a family formation effect resulting from migration, for migrants to England and Wales in the 1991-2001 period. An interesting by-product is the finding that there is an important age and parity relationship for teenagers. Looking ahead, there will be many opportunities to build on the research in this thesis and further estimate the fertility of migrants in the 2001-2011 period with the inclusion of 2011 census data.

References

Adelstein, A. M. (1976) Policies of Office of Population Censuses and Surveys – Philosophy and Constraints, *British Journal of Preventive & Social Medicine*, **30**, 1, 1-10.

Andersson, G. (2004) Childbearing after migration: Fertility patterns of foreign-born women in Sweden, *International Migration Review*, **38,** 2, 747-774.

Allison, P. D. (1982) Discrete-Time Methods for the Analysis of Event Histories, *Sociological Methodology*, 13: 61-98.

Arai, L. (2007) Peer and neighbourhood influences on teenage pregnancy and fertility: Qualitative findings from research in English communities, *Health & Place,* 13(1): 87-98.

Babb, P. and Bethune, A. (1995) Trends in births outside marriage, *Population Trends*, 81, 17-22.

Babb, P. and Hattersley, L. (1992) *An examination of the quality of OPCS Longitudinal Study data for use in fertility analysis*, LS Demographic Analysis Section, OPCS, London.

Blackwell, L., Lynch, K., Smith, J., and Goldblatt, P. (2003) *Longitudinal Study 1971-2001: Completeness of Census Linkage, Series LS no. 10.*, Office for National Statistics, London.

Blackwell, L., Akinwale, B., Antonatos, A., and Haskey, J. (2005) Opportunities for new research using the post-2001 ONS Longitudinal Study, *Population Trends*, **121**, 8-16.

Blossfeld, H.-P. and Huinink, J. (1991) Human capital investments or norms of role transition? How women's schooling and career affect the process of family formation. *American Journal of Sociology*, 97(1): 143-168.

Berrington, A. (2004) Perpetual postponers? Women's, men's and couple's fertility and subsequent fertility behaviour, *Population Trends,* **117**, 9-19.

Boseley, S. and Saner, E. (2009) *The new baby boom*, [on-line] The Guardian, London, http://www.guardian.co.uk/theguardian/2009/aug/29/new-baby-boom-boston-fertility Accessed 09-07-2010.

Buxton, J, Clarke, L, Grundy, E, and Marshall, C. E. (2005) The long shadow of childhood: associations between parental social class and own social class, educational attainment and timing of first birth; results from the ONS Longitudinal Study, *Population Trends*, 121, 17-26.

Coleman, D. (1994) Trends in fertility and intermarriage among immigrant populations in Western Europe as measures of integration, *Journal of Biosocial Science,* **26**, 1, 107-136.

Coleman, D. (2007) Demographic diversity and the ethnic consequences of immigration - key issues that the Commission's report left out, *Vienna Yearbook of Population Research 2007,* 5-12.

Coleman, D. A. and Dubuc, S. (2010) The fertility of ethnic minorities in the UK, 1960-2006, *Population Studies,* 64(1): 19-41.

Coombes, M., Champion, T., and Raybould, S. (2007) Did the Early A8 In-migrants to England go to Areas of Labour Shortage? *Local Economy,* **22,** 4, 335 – 348.

Dugard, P., Todman, J., and Staines, H. (2010) *Approaching Multivariate Analysis – A practical introduction,* Second Edition, Taylor & Francis, London.

Dunnell, K. (2007) The Changing Demographic Picture of the UK – National Statistician's Annual Article on the Population, *Population Trends*, **130,** 9-21.

Dustmann, C. and Fabbri, F. (2005) Immigrants in the British Labour Market, *Fiscal Studies,* 26(4): 423-470.

Ekert-Jaffé, O., Joshi, H., Lynch, K., Mougin, R., Rendall, M., and Shapiro, D. (2002) Fertility, Timing of Births and Socio-economic Status in France and Britain: Social Policies and Occupational Polarization. *Population,* **53,** 3, 475-507.

Farr, W. (1839) 2nd Letter to the Registrar General published in the Annual Report of the Registrar General.

Ford, K. (1990) Duration of Residence in the United-States and the Fertility of United-States Immigrants, *International Migration Review,* **24,** 1, 34-68.

Foreign and Commonwealth Office, n.d. Registering a birth, [on-line] Foreign and Commonwealth Office, http://www.fco.gov.uk/en/travel-and-living-abroad/births-deaths-marriagescivil/registering-a-birth Accessed 2 February 2010.

Garssen, J. and Nicolaas, H. (2008) Fertility of Turkish and Moroccan women in the Netherlands: Adjustment to native level within one generation, *Demographic Research*, **19**, 33, 1249-1280.

Goldstein, S. and Goldstein, A. (1981) The Impact of Migration on Fertility - an Own Children Analysis for Thailand, *Population Studies-a Journal of Demography*, **35**, 2, 265-284.

Grundy, E. (1986) Migration and fertility behaviour in England and Wales – a record linkage study, *Journal of Biosocial Science*, 18, 4, 403-423.

Hattersley, L. (1999) *LS User Guide 18 – International migration data in the Longitudinal Study*, Office for National Statistics, LS Unit, London.

Hattersley, L. and Creeser, R. (1995) *Longitudinal Study 1971-1991 - History, organisation and quality of data*, HMSO, London.

Héran, F. and Pison, G. (2007) Two children per woman in France in 2006: are immigrants to blame? *Population & Societies,* **432,** 1-4.

Hinde, A. (1998) *Demographic methods*, Arnold, London.

Jefferies, J. (2001) A reluctance to embrace the one-child family in Britain? Paper presented at Bad Herrenalb. Paper prepared for EURESCO conference, 'The Second Demographic Transition in Europe', Bad Herrenalb, Germany, 23-28 June, 2001. Bad Herrenalb, Germany. 23-28 June, 2001.

Kneale, D. and Joshi, H. (2008) Postponement and childlessness: Evidence from two British cohorts, *Demographic Research,* 19: 1935-1968.

Kulu, H. (2003) Migration and Fertility: Competing Hypotheses Re-examined, *MPIDR Working Paper WP 2003-035*, working Paper from the Max Planck Institute for Demographic Research.

Matheson, J. (2009) National Statistician's Annual Article on the Population: a Demographic Review, *Population Trends*, **138,** 7-21.

Matz, D., Hill, R. and Heath, T. (2001) *Asylum Statistics United Kingdom 2000,* 17/01, 25 September 2001, Home Office, London.

Murphy, M. (2009) Where have all the children gone? Women's reports of more childlessness at older ages than when they were younger in a large-scale continuous household survey in Britain, *Population Studies*, **63**, 2: 115 – 133.

Mussino, E. and Van Raalte, A. (2008) Fertility of Migrants: A comparative study between Italy and Russia, *MPIDR Working Paper WP 2008-026*, working Paper from the Max Planck Institute for Demographic Research.

Mussino, E. and Strozza, S. (2012) The fertility of immigrants after arrival: The Italian case, *Demographic Research*, **26**, 4, 99-130.

Mussino, E, Iaccarino, C., Prati, S., and Strozza, S. (2009) Short-term reproductive behaviour of foreign women who became mothers between 2002-2006 in Italy, *Vienna Yearbook of Population Research*, 63-82.

Ní Bhrolcháin, M. (1988) Contraceptive confidence idea: An empirical investigation, *Population Studies*, **42**, 2, 205-225.

Ní Bhrolcháin, M. and Beaujouan, E. (2012) Fertility postponement is largely due to rising educational enrolment, *Population Studies,* 66(3): 311-327.

Ní Bhrolcháin, M., Beaujouan, E., and Murphy, M. (2011) Reported childlessness: a further look at cohort estimates based on survey time-series data, *Centre for Population Change Working Paper*, Number 11.

O'Leary, L., Natamba, E., Jefferies, J. and Wilson, B. (2010) Fertility and partnership status in the last two decades, *Population Trends,* 140: 5-35.

Office for National Statistics (2007) *Both UK and foreign-born women contribute to rise in fertility* [on-line] Office for National Statistics, Newport, UK, http://www.statistics.gov.uk/pdfdir/fertility1207.pdf Accessed 09-07-2010.

Office for National Statistics, 'no date' a. *Linking 2001 Census data to the ONS Longitudinal Study*, Office for National Statistics, [online], http://www.ons.gov.uk/about/who-we-are/our-services/longitudinal-study/longitudinalstudy-history-and-processes/linking-census-data-to-the-ls/more-about-the-2001-linking-exercise.pdf Accessed 7 May 2009.

Office for National Statistics, n.d.a. *ONS Longitudinal Study - Quality of tracing at the 2001 Census*, [on-line], Office for National Statistics, http://www.ons.gov.uk/about/who-we-are/ourservices/longitudinal-study/data-quality/quality-of-tracing/quality-of-tracing-at-the-2001-census–112kb–pdf-.pdf Accessed 7 May 2009.

Office for National Statistics, n.d.b. *ONS Longitudinal Study - Sampling quality at the 1991 Census,* [on-line], Office for National Statistics, http://www.ons.gov.uk/about/who-we-are/ourservices/longitudinal-study/data-quality/sampling-quality/1991-census-sampling-fractions–45kb-pdf-.pdf Accessed 7 May 2009.

Office for National Statistics, n.d.c. *Topic guide to: Live Births and Stillbirths*, [on-line] Office for National Statistics, http://www.statistics.gov.uk/hub/population/births-and-fertility/livebirths-and-stillbirths Accessed 2 February 2010.

Office for National Statistics, n.d.d. *Linking census data to the LS,* [on-line] Office for National Statistics, http://www.ons.gov.uk/about/who-we-are/our-services/longitudinal-study/longitudinalstudy-history-and-processes/linking-census-data-to-the-l Accessed 2 February 2010.

Office for National Statistics, n.d.e. *New births into the Longitudinal Study*, [on-line] Office for National Statistics, http://www.ons.gov.uk/about/who-we-are/our-services/longitudinalstudy/data-quality/event-sampling-and-linkage/new-births-into-the-ls/new-births-by-sex-1971-2005.pdf Accessed September 2009.

Office for National Statistics, n.d.f. *Updating the LS data with events data,* [on-line] Office for National Statistics, http://www.ons.gov.uk/about/who-we-are/our-services/longitudinalstudy/longitudinal-study-history-and-processes/updating-the-ls-data-with-events-data/index.html Accessed March 2010.

Office for National Statistics, n.d.g. *Annual Population Survey (APS),* [on-line] Office for National Statistics, http://www.ons.gov.uk/ons/about-ons/who-we-are/services/unpublished-data/social-survey-data/aps/index.html Accessed September 2012.

Office for National Statistics (2011) *Feasibility of linking Migrant Workers Scan data and GP Patient Register data,* Confidential internal report not for publication or circulation, Office for National Statistics Centre for Demography.

Office for National Statistics (27 March 2012) *Uncertainty in local authority mid-year population estimates - Office for National Statistics Research Report*, [on-line] Office for National Statistics, http://www.ons.gov.uk/ons/guide-method/method-quality/imps/latest-news/uncertainty-in-la-mypes/index.html Accessed September 2012.

Office of Population Censuses and Surveys (OPCS) (1973) *Cohort Studies: New Developments, Studies on Medical and Population Subjects*, No. 25, London.

Peach, C. (2006) South Asian migration and settlement in Great Britain, 1951–2001, *Contemporary South Asia*, **15**, 2, 133–146.

Portanti, M. and Whitworth, S. (2009) A comparison of the characteristics of childless women and mothers in the ONS Longitudinal Study, *Population Trends*, **136**: 10-20.

Raymer, J. and Rogers, A. (2008) Applying model migration schedules to represent age-specific migration flows. In, Raymer, J. and Willekens, F. (eds.) *International Migration in Europe: Data, Models and Estimates,* Wiley, Chichester, UK, 175-192.

Rendall, M. (2003) How important are intergenerational cycles of teenage motherhood in England and Wales? A comparison with France, *Population Trends*, 111, 27-37.

Rendall, M. and Ball, D. J. (2004) Immigration, emigration and the ageing of the overseas-born population in the United Kingdom, *Population Trends,* 116: 18-27.

Rendall, M. and Salt, J. (2005) The foreign-born population (Chapter 8). In Chappell, R. (ed) *Office for National Statistics, Focus on: People and Migration,* Palgrave Macmillan, 131-152.

Rendall, M. and Smallwood, S. (2003) Higher qualifications, first-birth timing and further childbearing in England and Wales, *Population Trends,* **111**, 18-25.

Rendall, M., Couet, C., Lappegard, T., Robert-Bobee, I., Ronsen, M., and Smallwood, S. (2005) First births by age and education in Britain, France and Norway, *Population Trends,* **121**, 27-34.

Rendall, M. S., Ekert-Jaffé, O., Joshi, H., Lynch, K., and Mougin, R. (2009) Universal versus Economically Polarized Change in Age at First Birth: A French-British Comparison, *Population and Development Review*, **35**, 1, 89-115.

Rogers, A., and Castro, L. J. (1981) *Model Migration Schedules*, International Institute for Applied System Analysis, Laxenburg, Austria.

Schoorl, J. J. (1990) Fertility Adaptation of Turkish and Moroccan Women in The Netherlands, *International Migration*, **28**, 4, 477-495.

Sharfman, A., Staples, V., and Hughes, H. (2010) Feasibility research on the potential use of Migrant Workers Scan data to improve migration and population statistics*, Population Trends,* 140, 36-52.

Sigle-Rushton, W. (2008) England and Wales: Stable fertility and pronounced social status differences, *Demographic Research*, **19,** 15, 455-502.

Singer, J. D. and Willett, J. B. (2003) *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence,* Oxford University Press, Oxford.

Smallwood, S. (2002) New estimates of trends in births by birth order in England and Wales, *Population Trends,* 108: 32-48.

Smallwood, S. and Lynch, K. (2010) An analysis of patient register data in the Longitudinal Study – what does it tell us about the quality of the data? *Population Trends*, 141, 151-169.

Sobotka, T. (2008) Overview Chapter 7: The rising importance of migrants for childbearing in Europe, *Demographic Research* **19,** 9, 225-248.

Sobotka, T. (2010) Fertility in Europe: Cloudy past and uncertain future. In *POPFEST 2010*. St. Andrews, Scotland 28-30th June 2010. BSPS: London.

Steele, F. (2005) *Event History Analysis*, Economic and Social Research Council National Centre for Research Methods, NCRM Methods Review Papers, NCRM/004, accessed 08/11/2012, http://eprints.ncrm.ac.uk/88/1/MethodsReviewPaperNCRM-004.pdf

Stephen, E. H. and Bean, F. D. (1992) Assimilation, Disruption and the Fertility of Mexican-Origin Women in the United States, *International Migration Review,* **26**, 1, 67-88.

Stevenson, J. (1973) *Cradle-to-grave tabs on 500,000 Britons*, The Daily Mail, London, Page 11.

Toulemon, L. (2004) Fertility among immigrant women: new data, a new approach, *Population & Societies*, No. **400**, 1-4.

Tromans, N., Natamba, E., Jefferies, J., and Norman, P. (2008) Have national trends in fertility between 1986 and 2006 occurred evenly across England and Wales? *Population Trends,* **133**, 7-19.

Tromans, N., Natamba, E., and Jefferies, J. (2009) Have women born outside the UK driven the rise in UK births since 2001? *Population Trends,* **136,** 28-42.

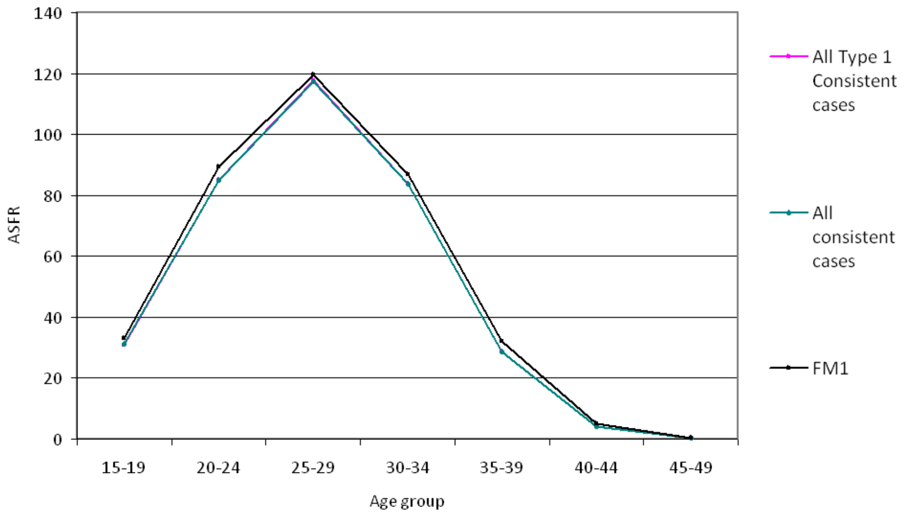
Werner, B. (1988) Spacing of births to women born in 1939-59: evidence from the OPCS Longitudinal Study, *Population Trends,* **52,** 20-25.

Appendix A

Additional figures from Chapter 5 'Calculating fertility rates using the Office for National Statistics Longitudinal Study'

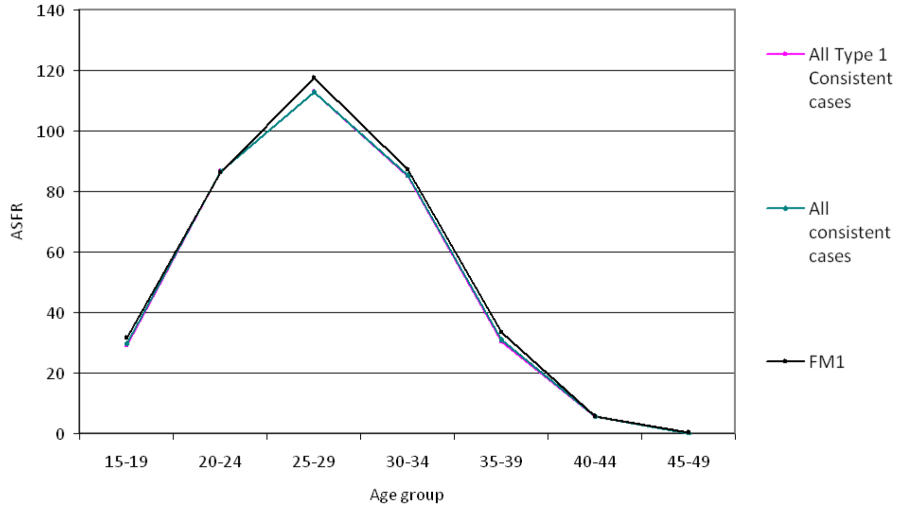
This appendix contains selected graphs and figures which were devised as part of the work for Chapter 5.

Figure A1: Consistent cases - age-specific fertility rates, 1991



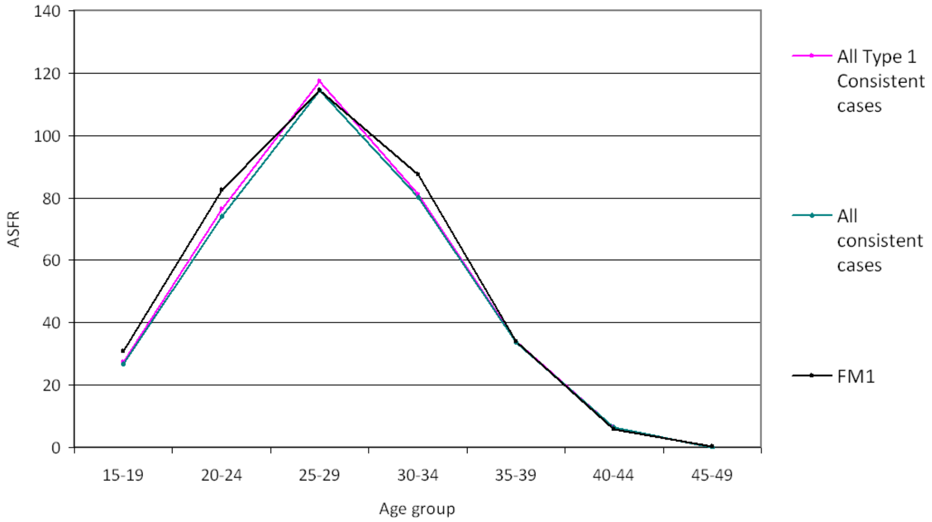
Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010.

Figure A2: Consistent cases - age-specific fertility rates, 1992



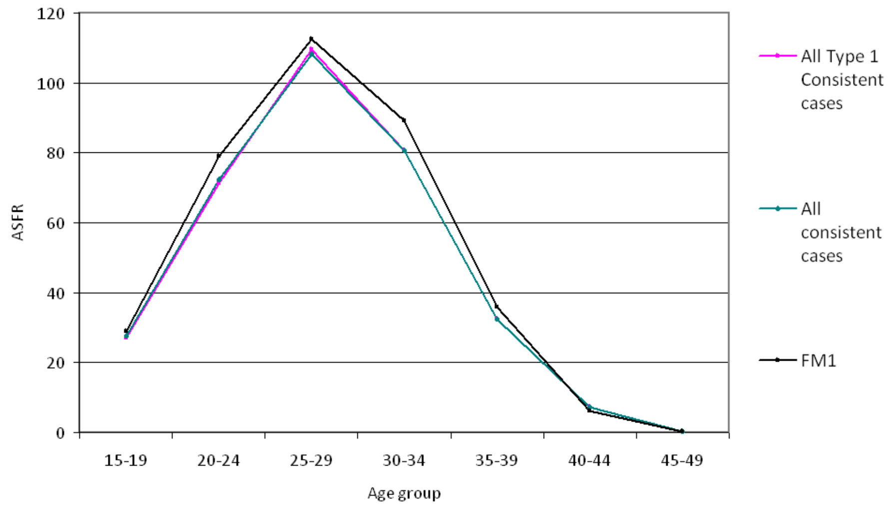
Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010.

Figure A3: Consistent cases - age-specific fertility rates, 1993



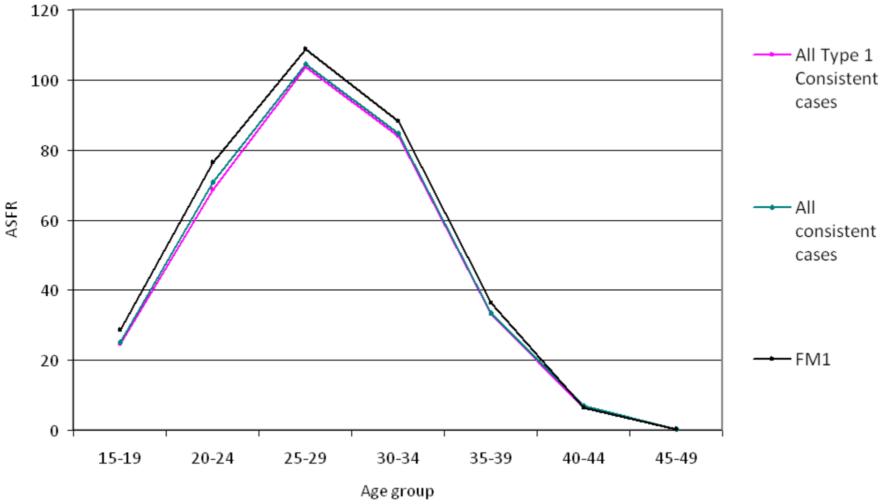
Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010.

Figure A4: Consistent cases - age-specific fertility rates, 1994



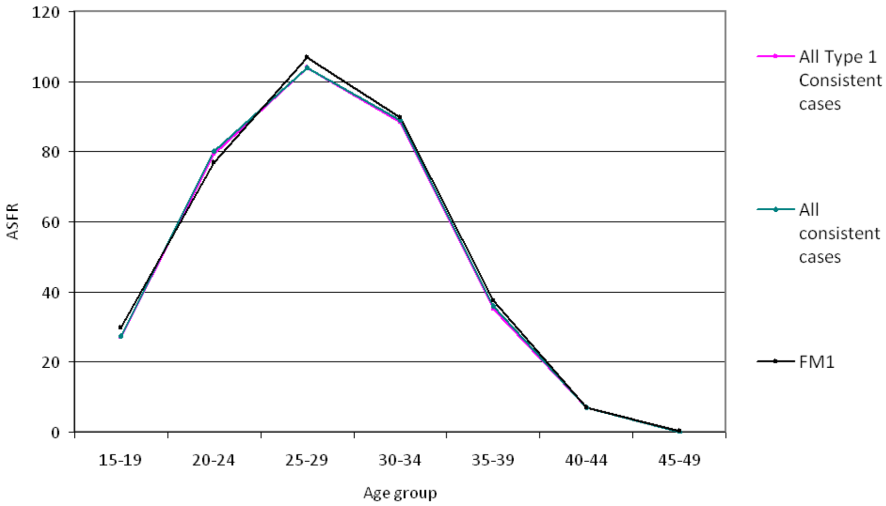
Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010.

Figure A5: Consistent cases - age-specific fertility rates, 1995



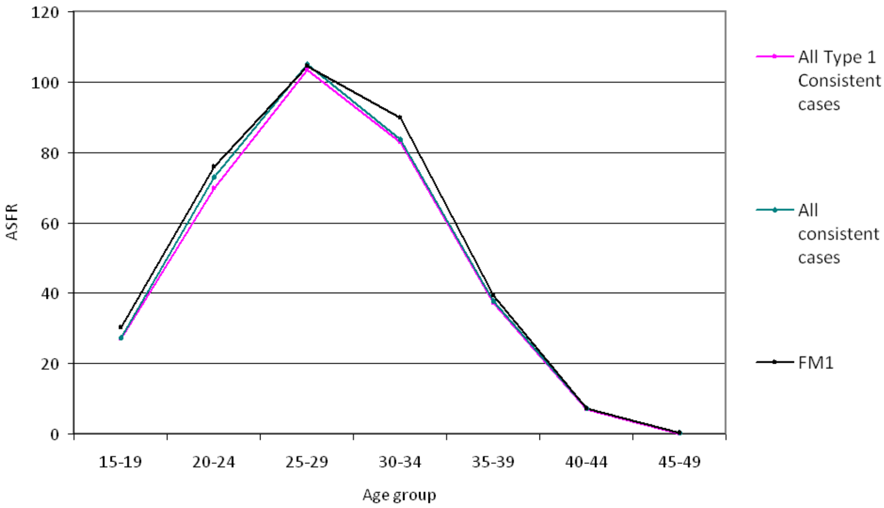
Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010.

Figure A6: Consistent cases - age-specific fertility rates, 1996



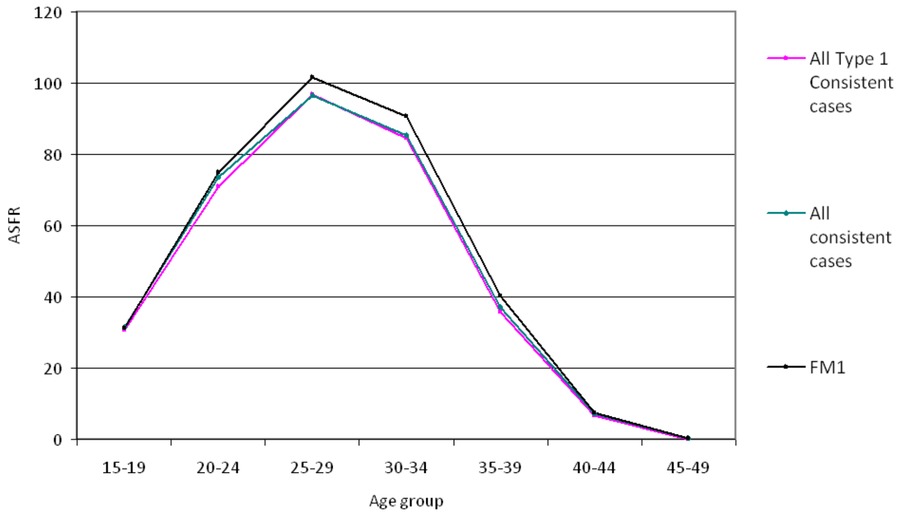
Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010.

Figure A7: Consistent cases - age-specific fertility rates, 1997



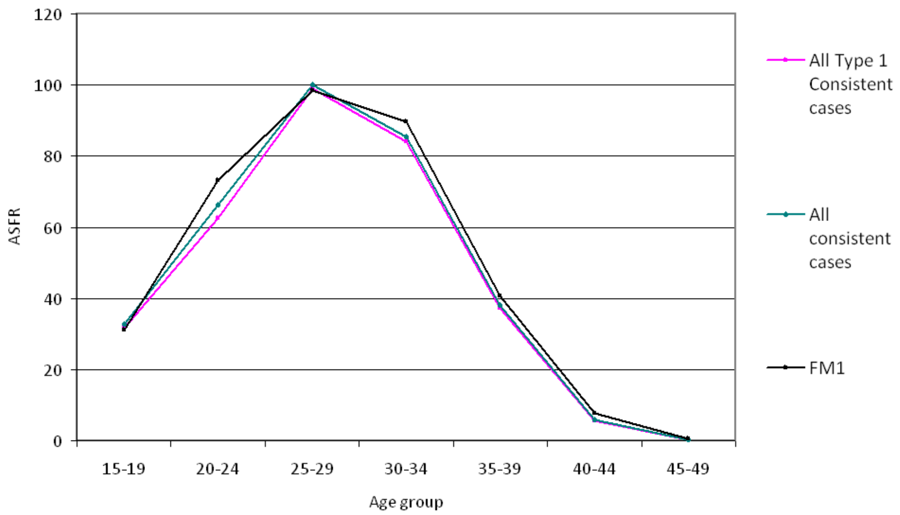
Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010.

Figure A8: Consistent cases - age-specific fertility rates, 1998



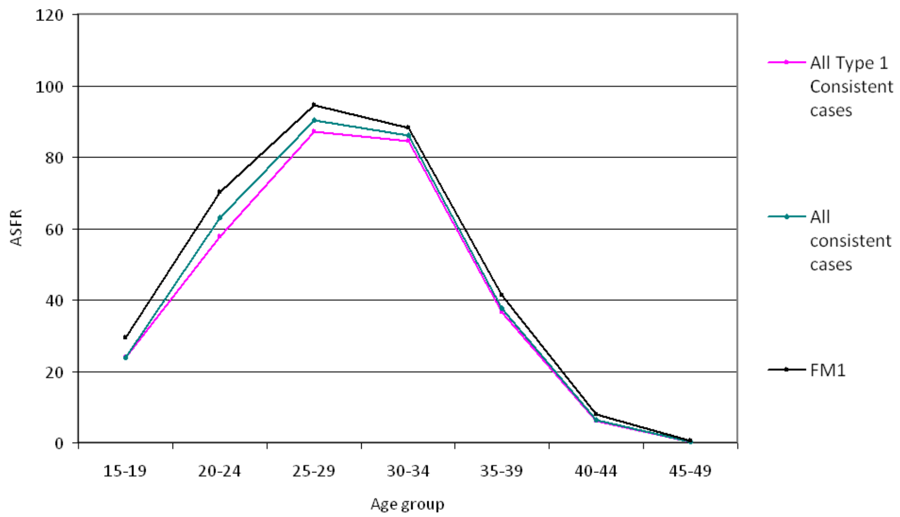
Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010.

Figure A9: Consistent cases - age-specific fertility rates, 1999



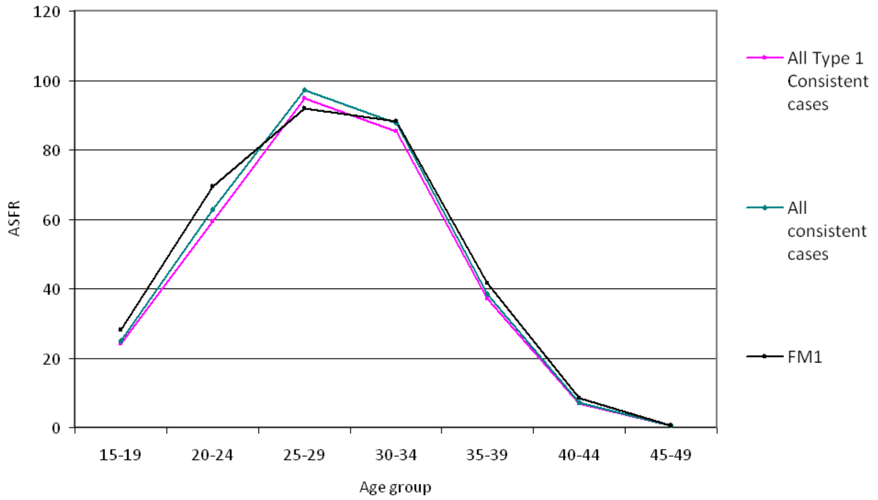
Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010.

Figure A10: Consistent cases - age-specific fertility rates, 2000



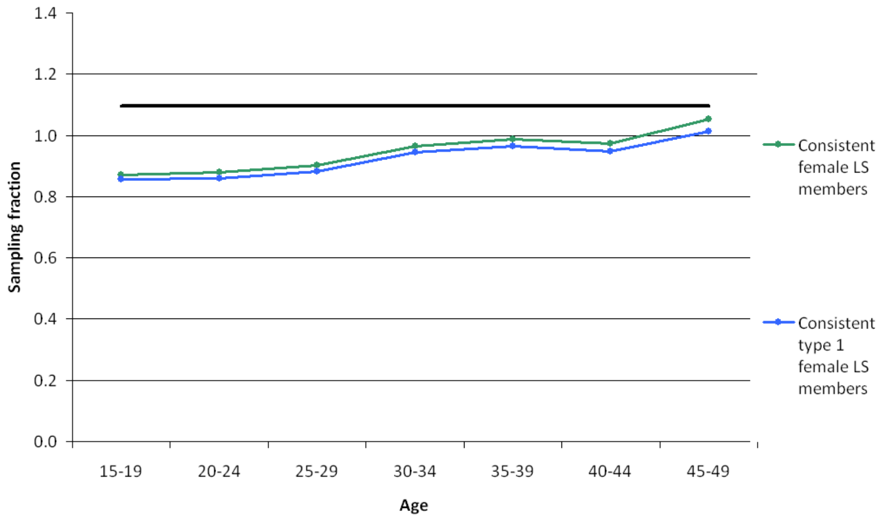
Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010.

Figure A11: Consistent cases - age-specific fertility rates, 2001



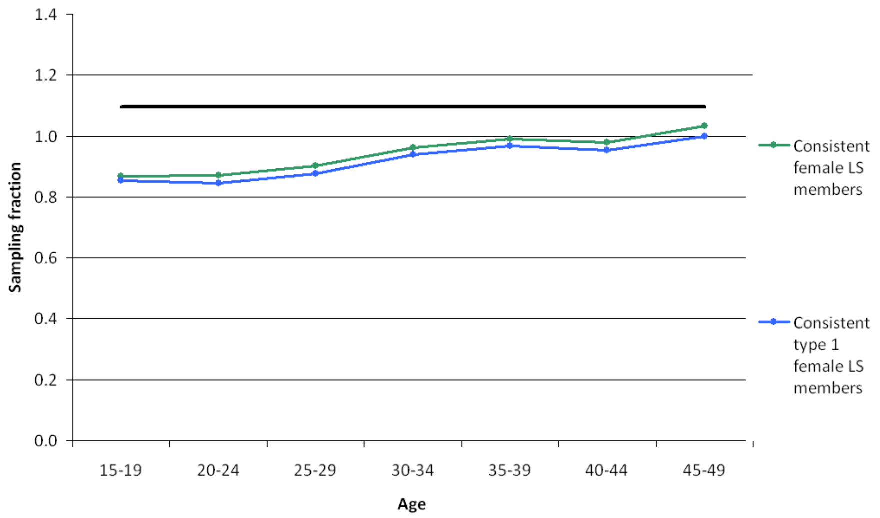
Own elaboration based on ONS LS, FM1 no.30, 2001, September 2010.

Figure A12: Consistent cases - sampling fraction of official statistics women by the LS, 1991



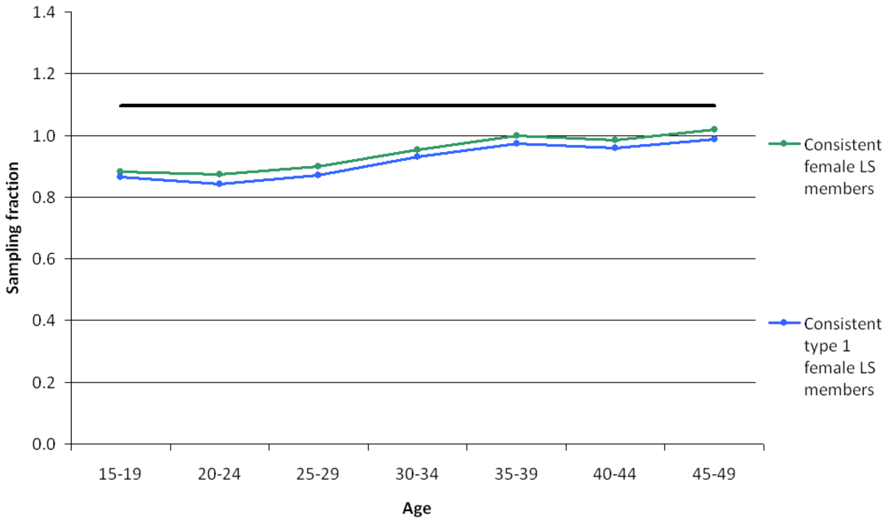
Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A13: Consistent cases - sampling fraction of official statistics women by the LS, 1992



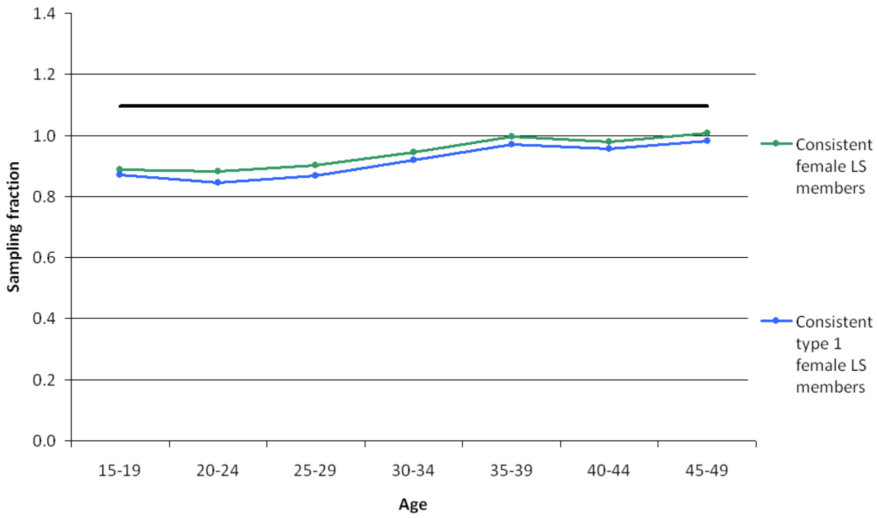
Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A14: Consistent cases - sampling fraction of official statistics women by the LS, 1993



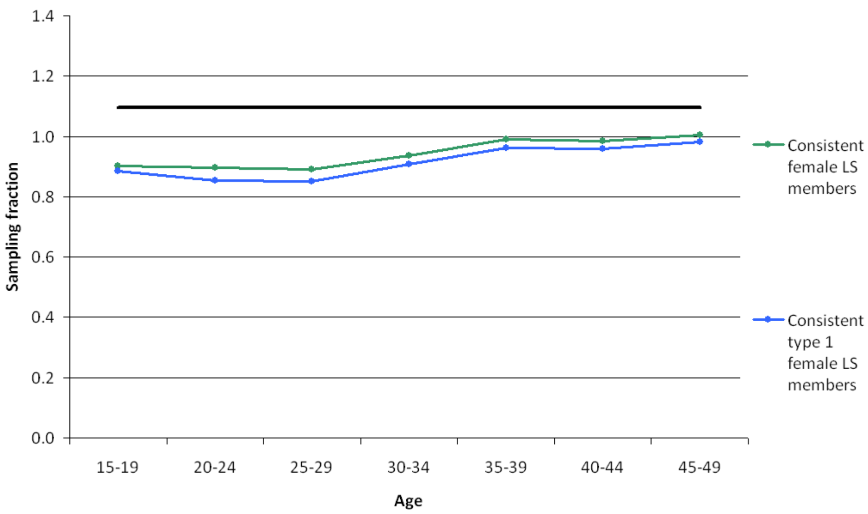
Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A15: Consistent cases - sampling fraction of official statistics women by the LS, 1994



Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

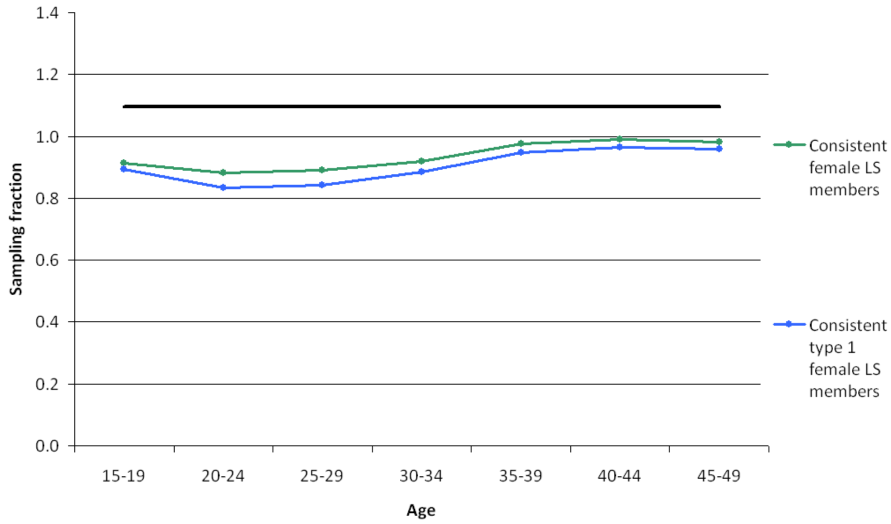
Figure A16: Consistent cases - sampling fraction of official statistics women by the LS, 1995



Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000

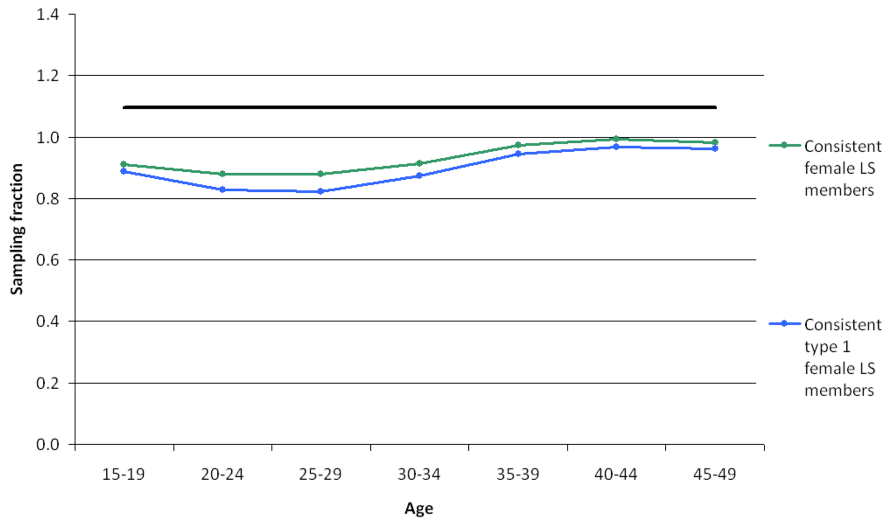
LA Population Studies: 07/10/04, June 2010.

Figure A17: Consistent cases - sampling fraction of official statistics women by the LS, 1996



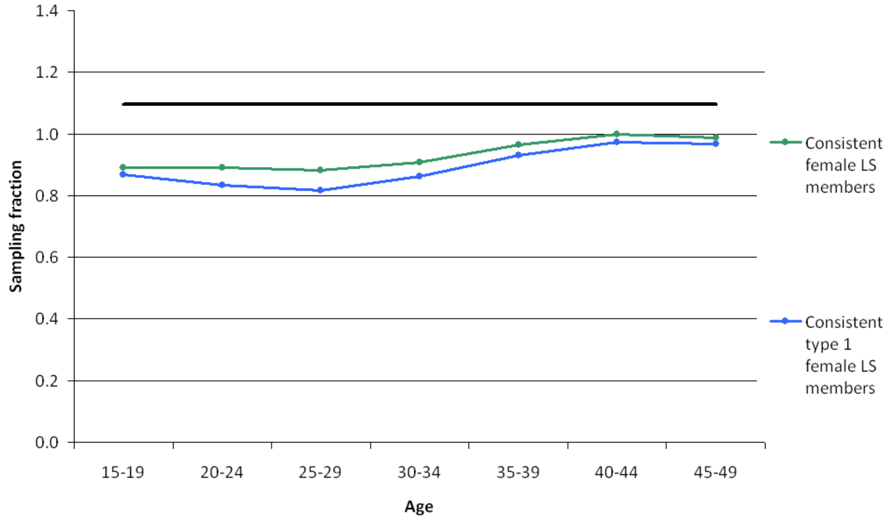
Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A18: Consistent cases - sampling fraction of official statistics women by the LS, 1997



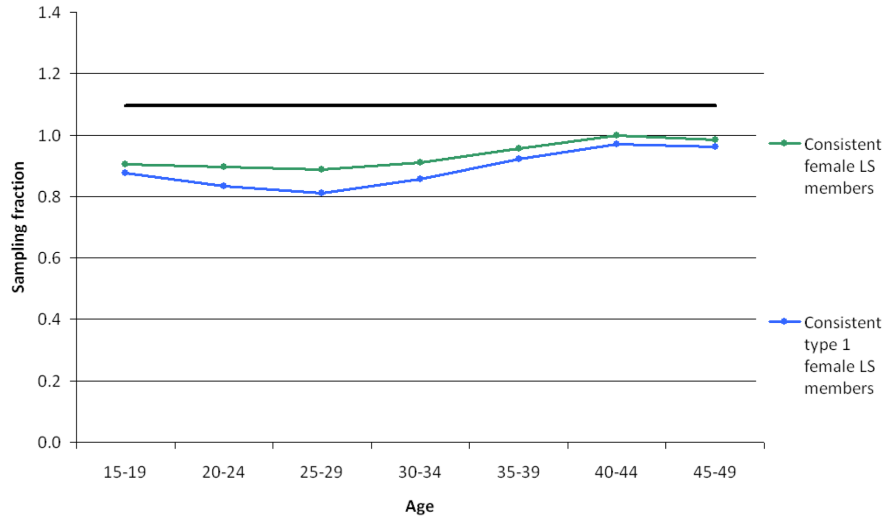
Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A19: Consistent cases - sampling fraction of official statistics women by the LS, 1998



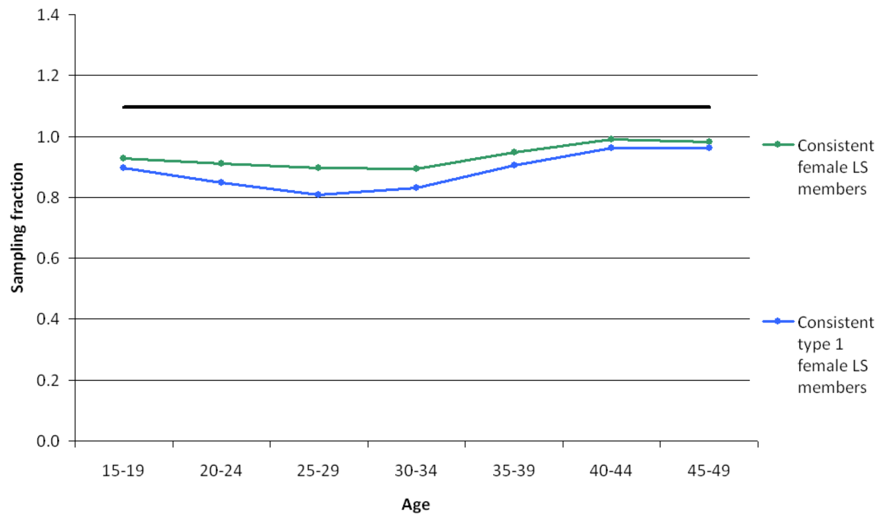
Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A20: Consistent cases - sampling fraction of official statistics women by the LS, 1999



Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

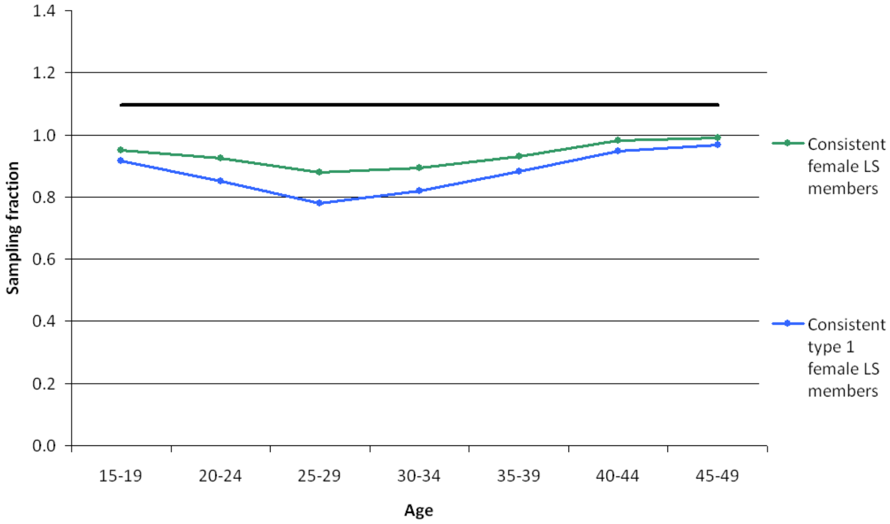
Figure A21: Consistent cases - sampling fraction of official statistics women by the LS, 2000



Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000

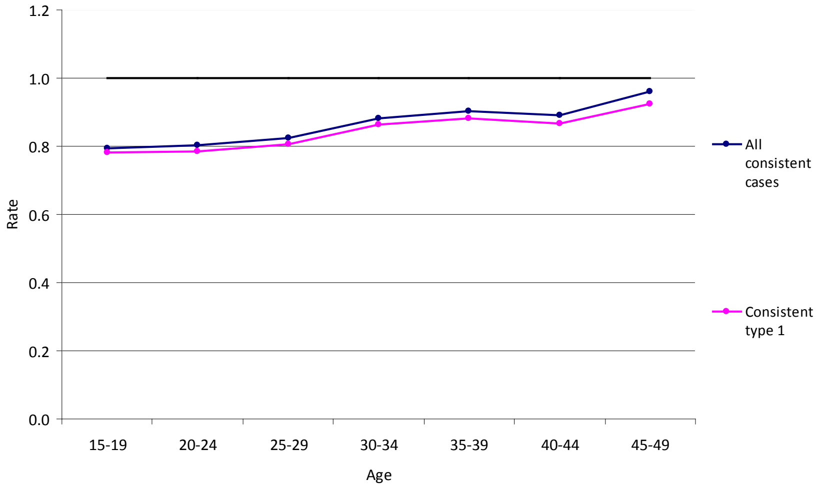
LA Population Studies: 07/10/04, June 2010.

Figure A22: Consistent cases - sampling fraction of official statistics women by the LS, 2001



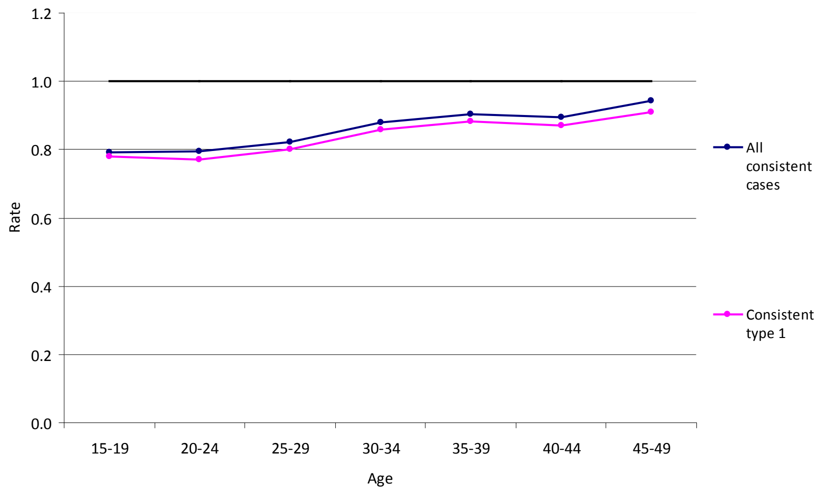
Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A23: Consistent cases - representation of LS women based on official statistics, 1991



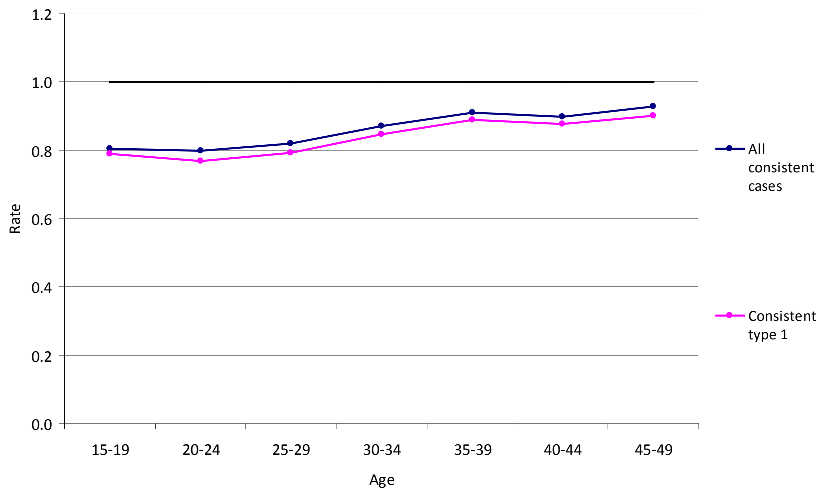
Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A24: Consistent cases - representation of LS women based on official statistics, 1992



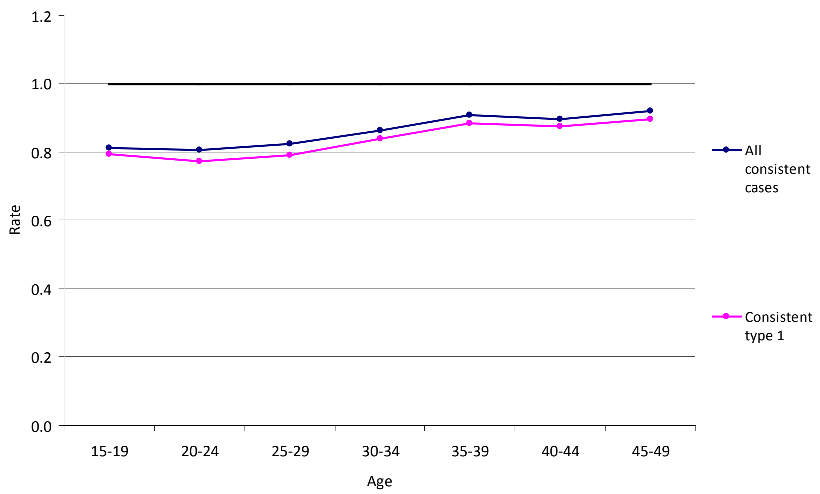
Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A25: Consistent cases - representation of LS women based on official statistics, 1993



Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

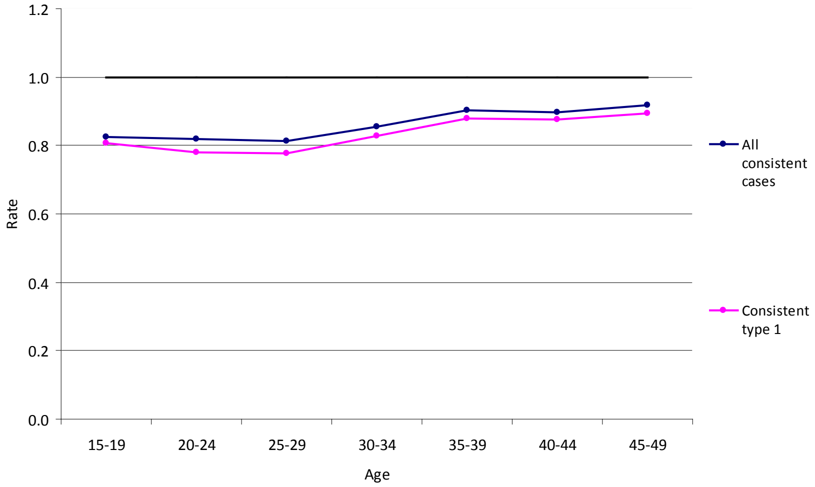
Figure A26: Consistent cases - representation of LS women based on official statistics, 1994



Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000

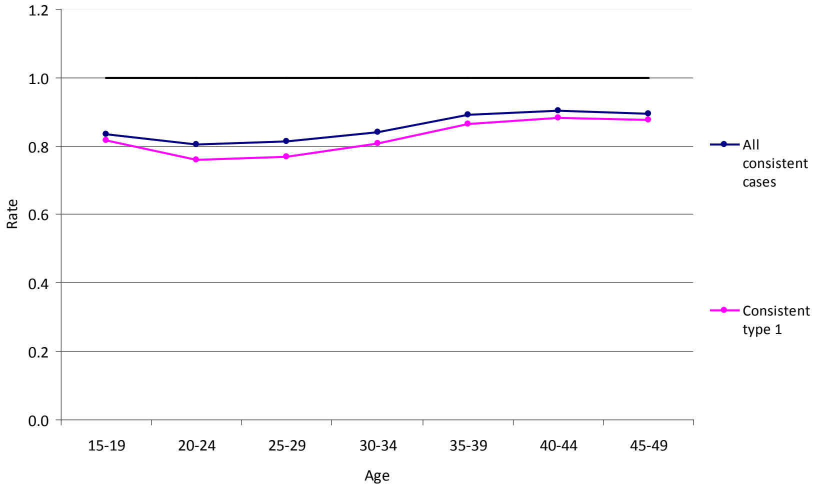
LA Population Studies: 07/10/04, June 2010.

Figure A27: Consistent cases - representation of LS women based on official statistics, 1995



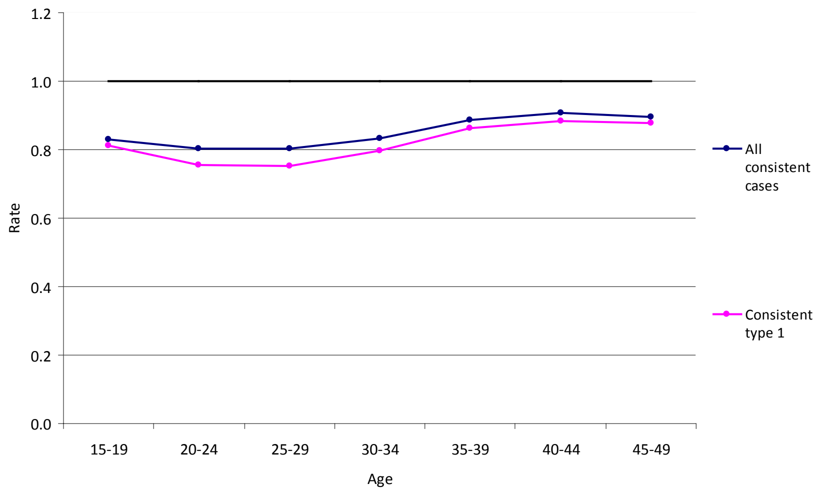
Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A28: Consistent cases - representation of LS women based on official statistics, 1996



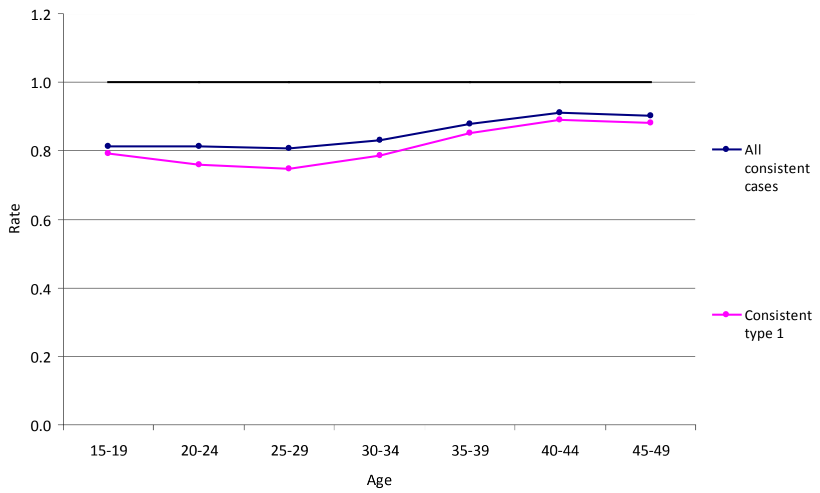
Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A29: Consistent cases - representation of LS women based on official statistics, 1997



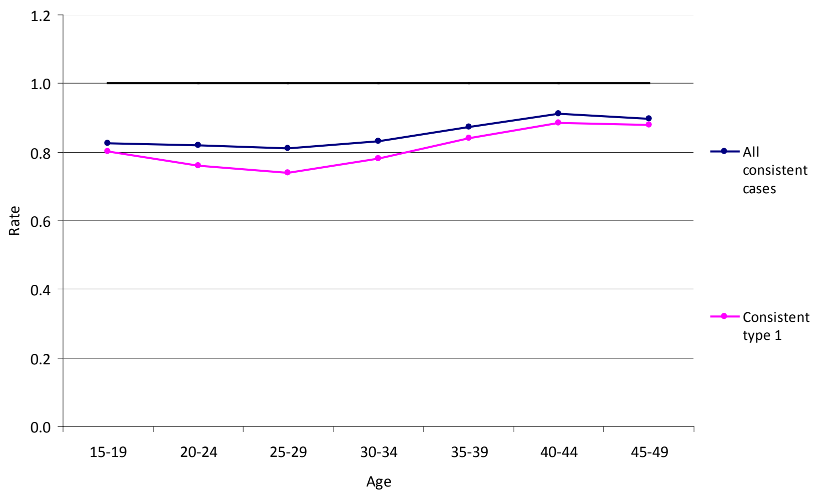
Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A30: Consistent cases - representation of LS women based on official statistics, 1998



Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

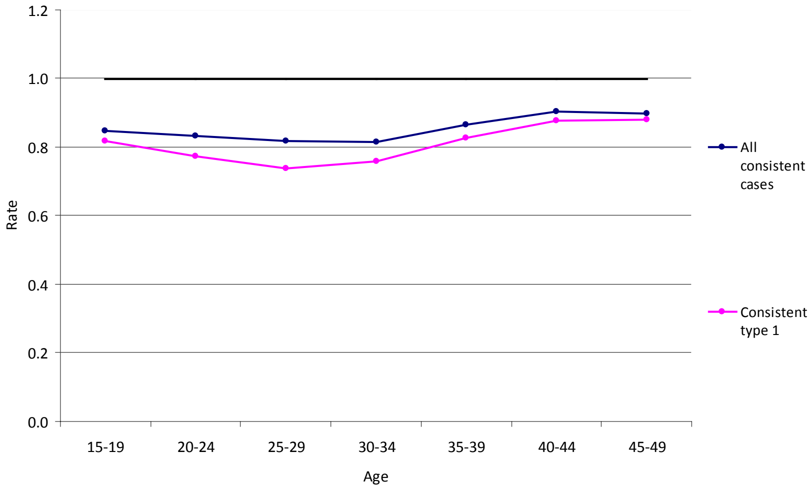
Figure A31: Consistent cases - representation of LS women based on official statistics, 1999



Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000

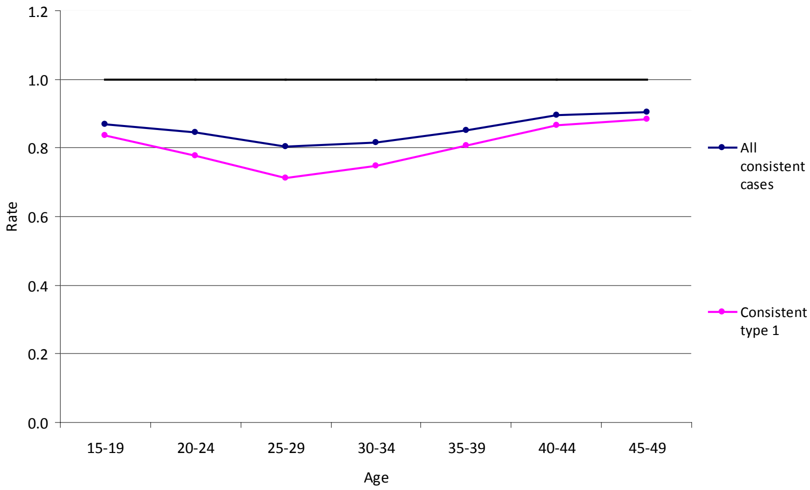
LA Population Studies: 07/10/04, June 2010.

Figure A32: Consistent cases - representation of LS women based on official statistics, 2000



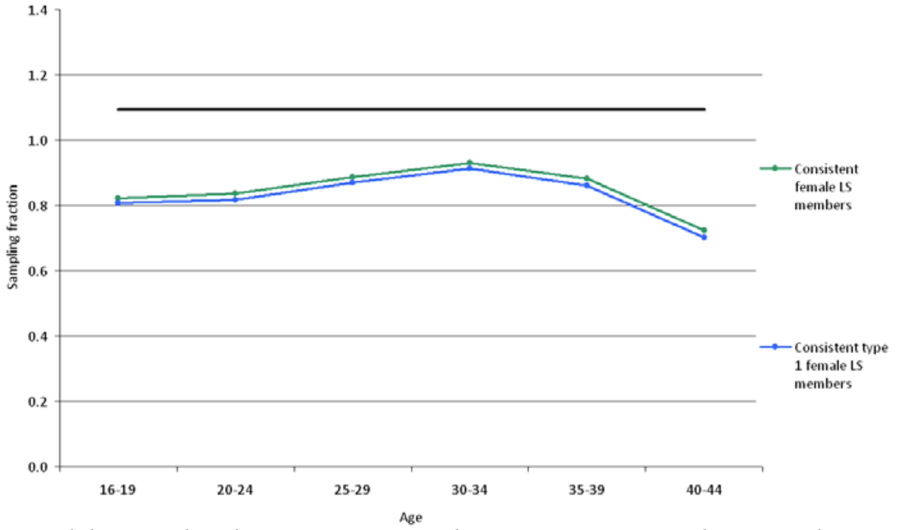
Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A33: Consistent cases - representation of LS women based on official statistics, 2001



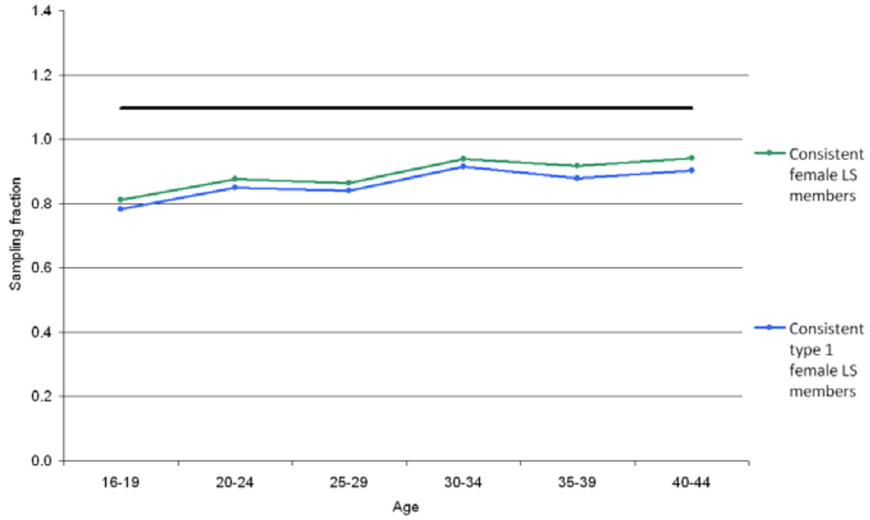
Own elaboration based on ONS LS, August 2010, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A34: Consistent cases - sampling fraction of official statistics births by the LS, 1991



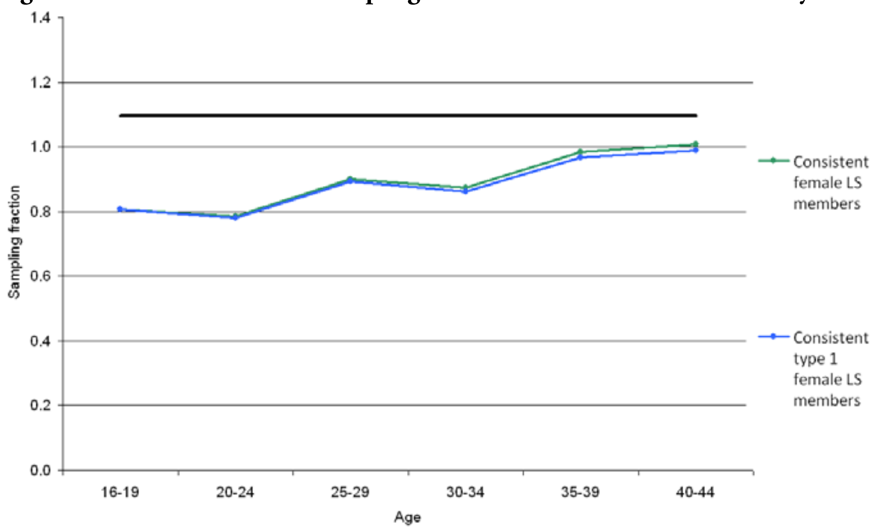
Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A35: Consistent cases - sampling fraction of official statistics births by the LS, 1992



Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

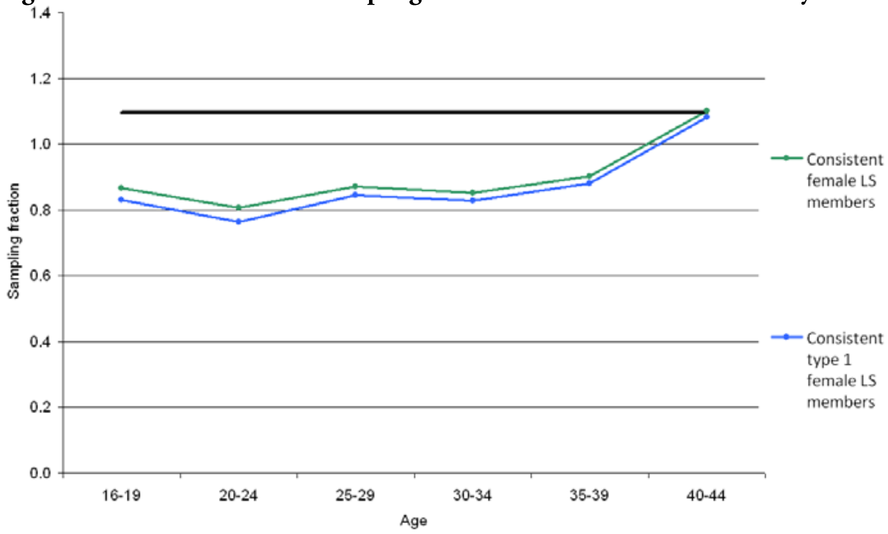
Figure A36: Consistent cases - sampling fraction of official statistics births by the LS, 1993



Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004,

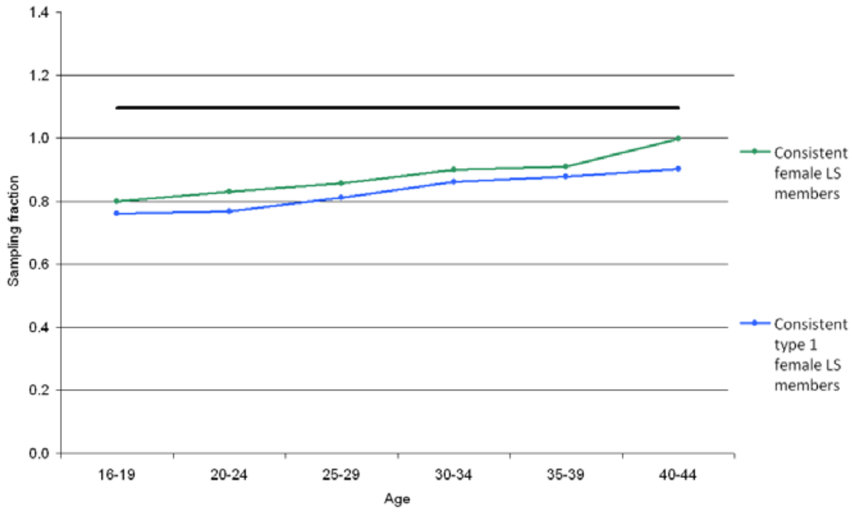
Maternities, Age of mother, a. all maternities, June 2010.

Figure A37: Consistent cases - sampling fraction of official statistics births by the LS, 1994



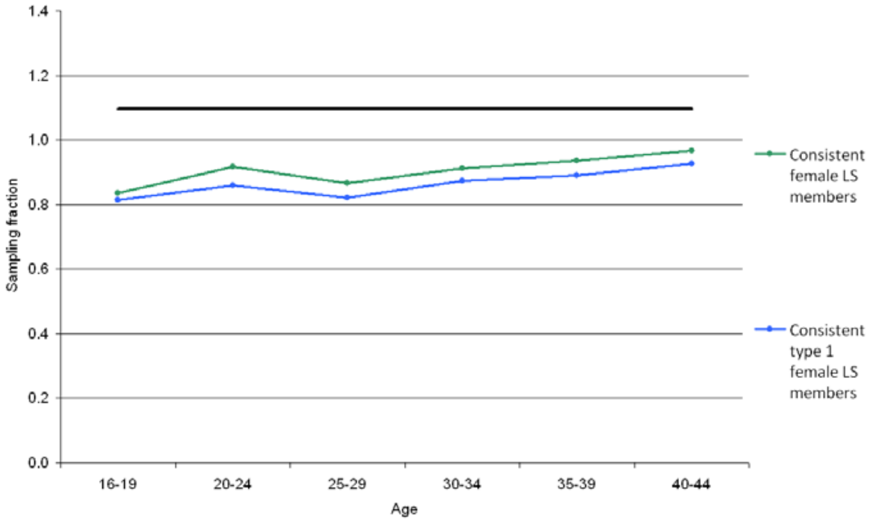
Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A38: Consistent cases - sampling fraction of official statistics births by the LS, 1995



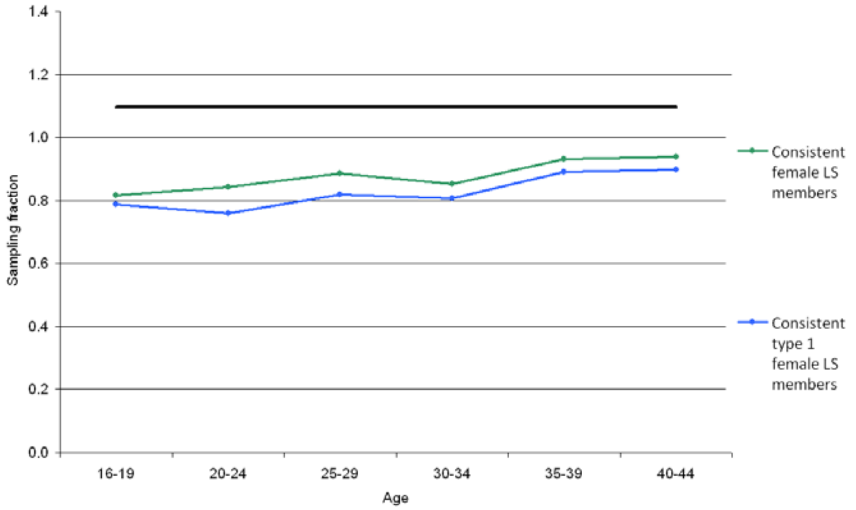
Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A39: Consistent cases - sampling fraction of official statistics births by the LS, 1996



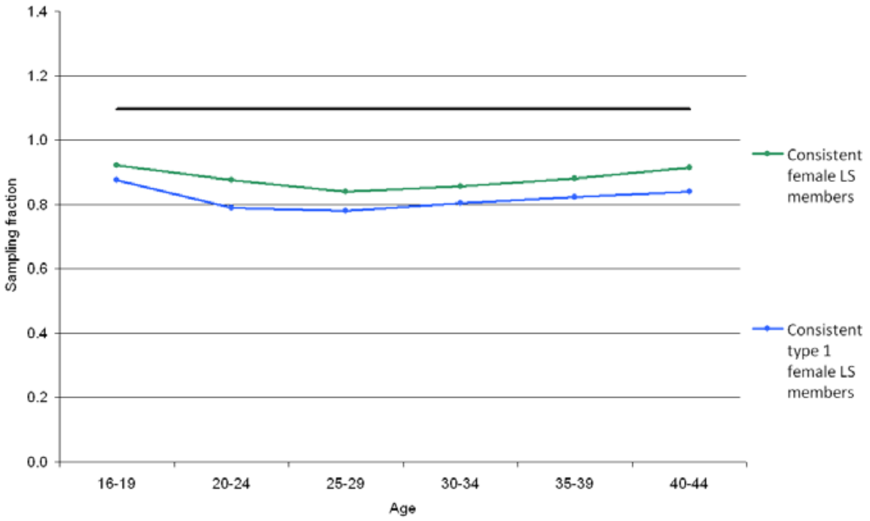
Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A40: Consistent cases - sampling fraction of official statistics births by the LS, 1997



Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

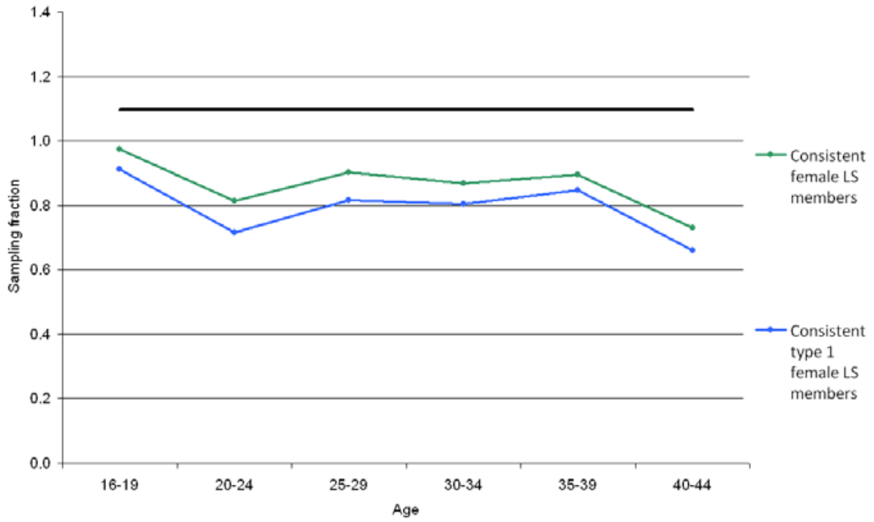
Figure A41: Consistent cases - sampling fraction of official statistics births by the LS, 1998



Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004,

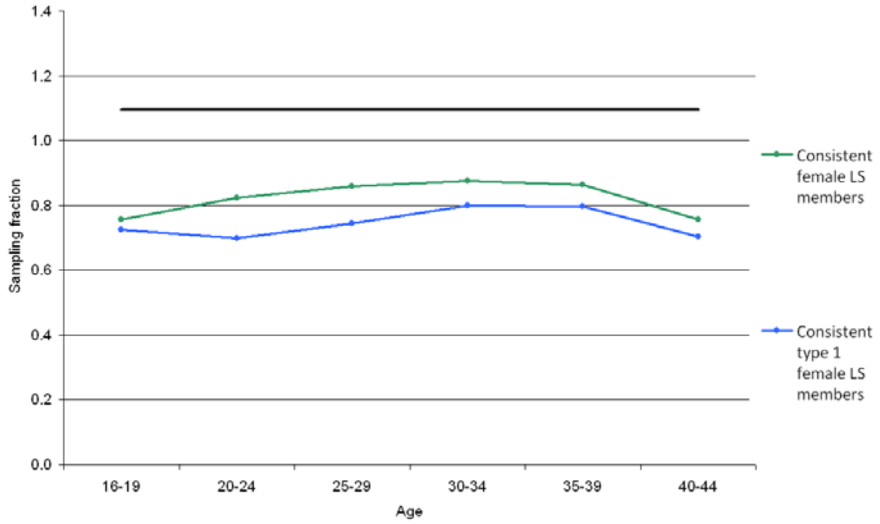
Maternities, Age of mother, a. all maternities, June 2010.

Figure A42: Consistent cases - sampling fraction of official statistics births by the LS, 1999



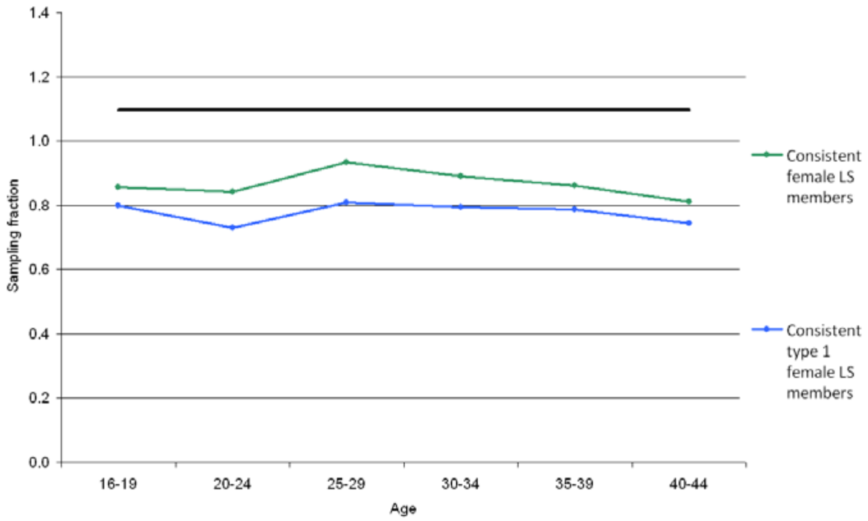
Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A43: Consistent cases - sampling fraction of official statistics births by the LS, 2000



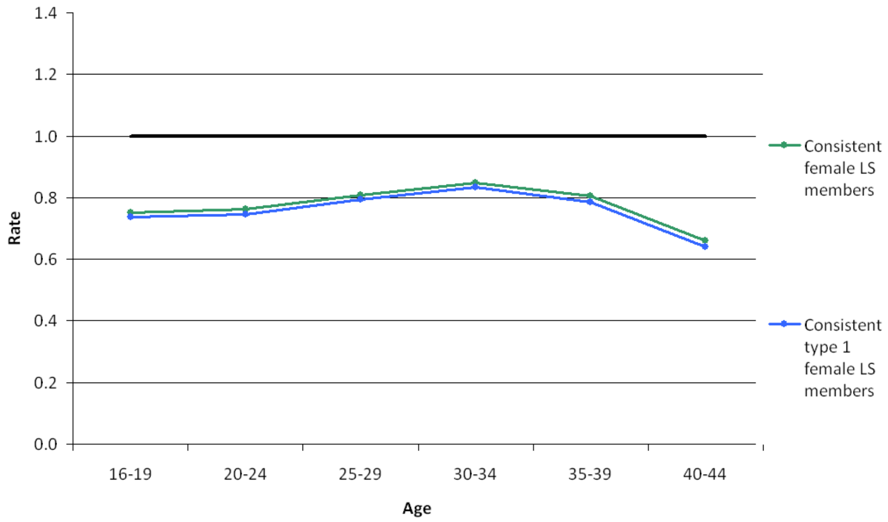
Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A44: Consistent cases - sampling fraction of official statistics births by the LS, 2001



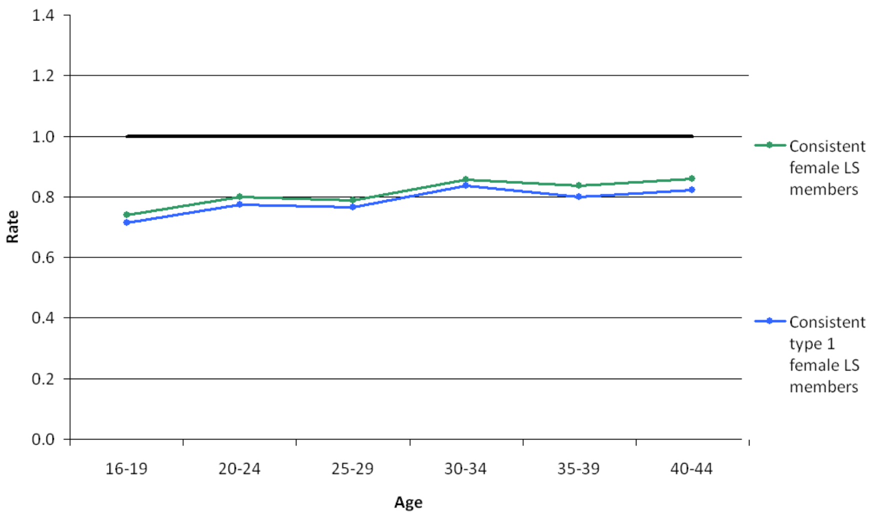
Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A45: Consistent cases - representation of LS births based on official statistics, 1991



Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

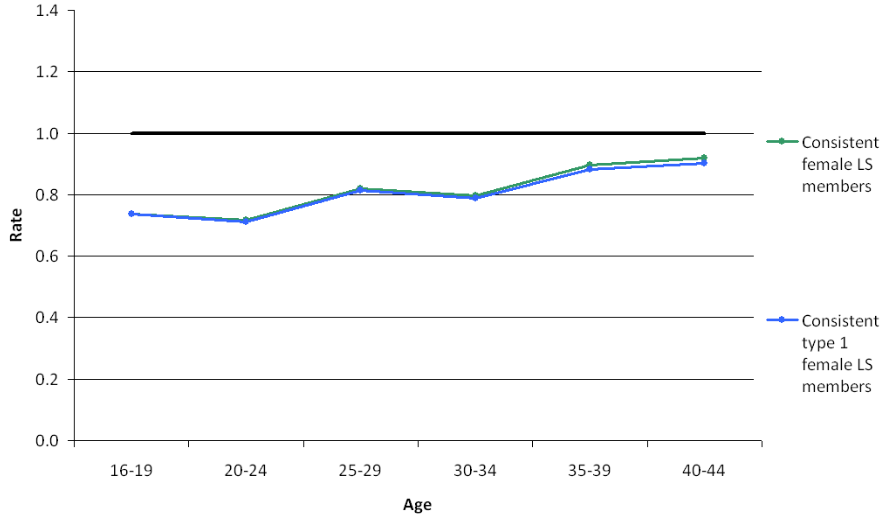
Figure A46: Consistent cases - representation of LS births based on official statistics, 1992



Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004,

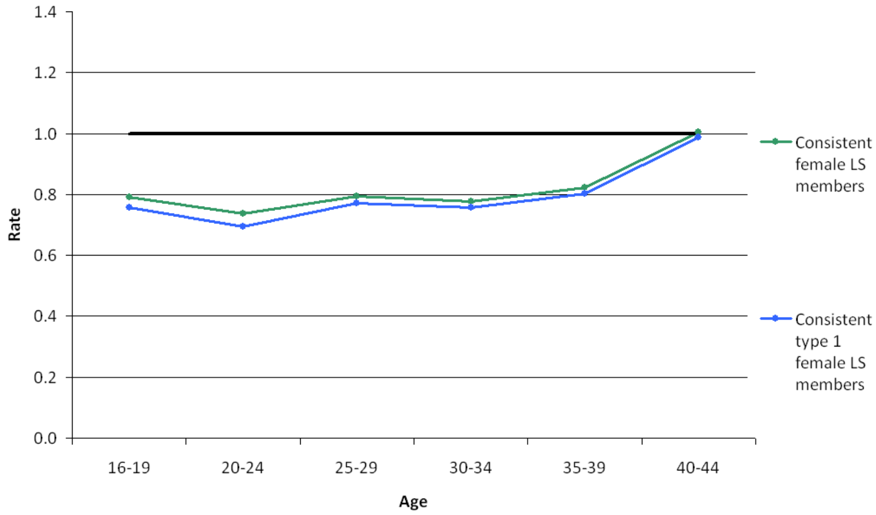
Maternities, Age of mother, a. all maternities, June 2010.

Figure A47: Consistent cases - representation of LS births based on official statistics, 1993



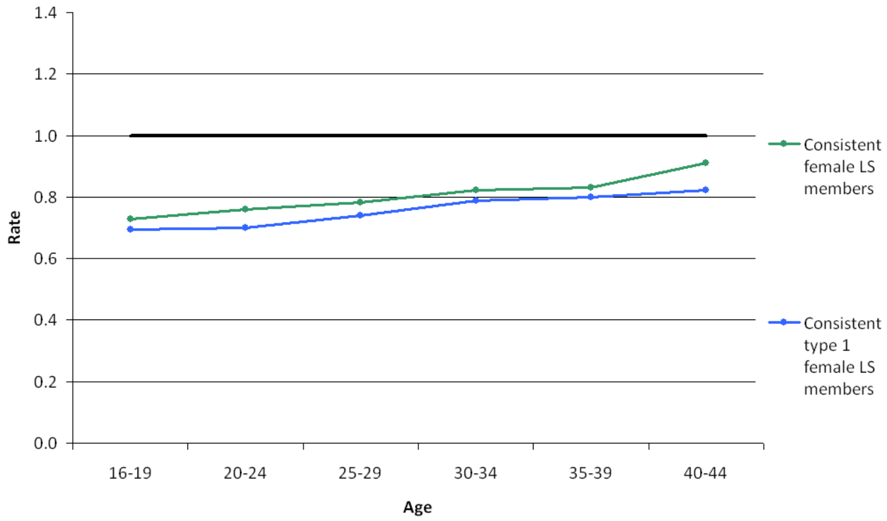
Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A48: Consistent cases - representation of LS births based on official statistics, 1994



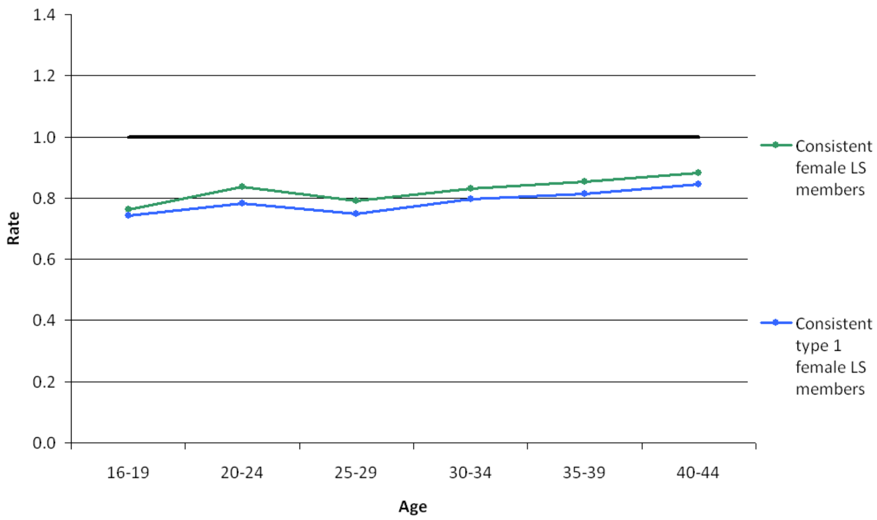
Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A49: Consistent cases - representation of LS births based on official statistics, 1995



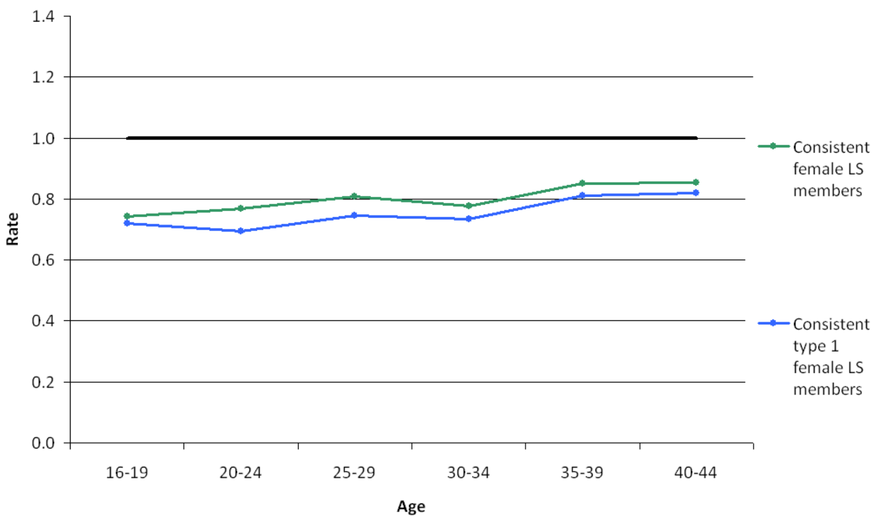
Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A50: Consistent cases - representation of LS births based on official statistics, 1996



Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

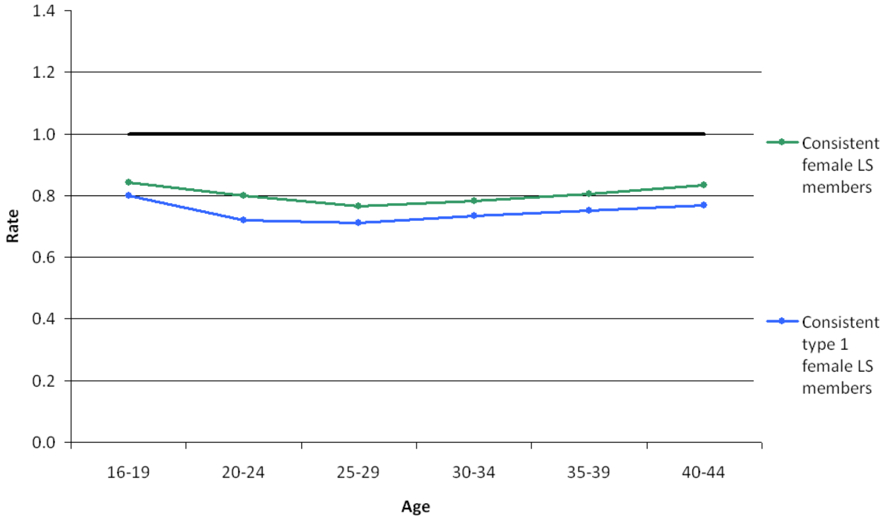
Figure A51: Consistent cases - representation of LS births based on official statistics, 1997



Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004,

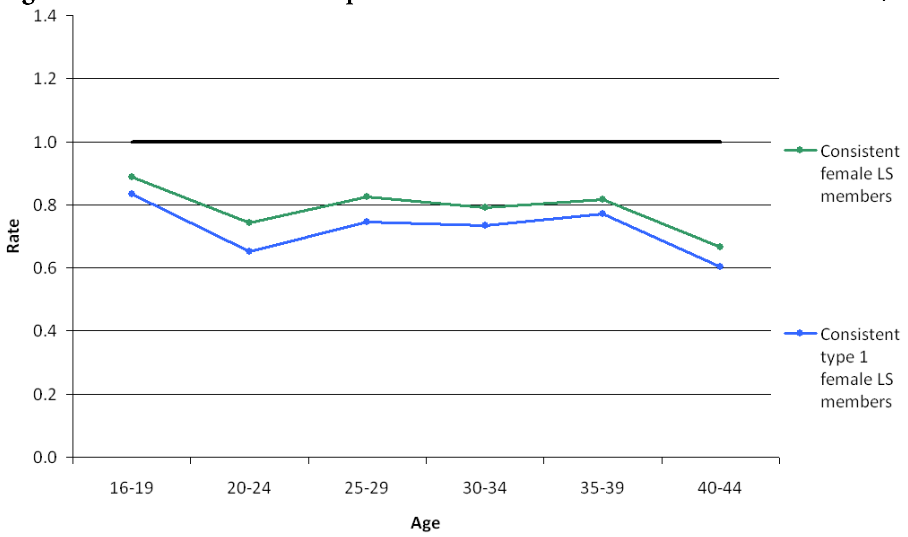
Maternities, Age of mother, a. all maternities, June 2010.

Figure A52: Consistent cases - representation of LS births based on official statistics, 1998



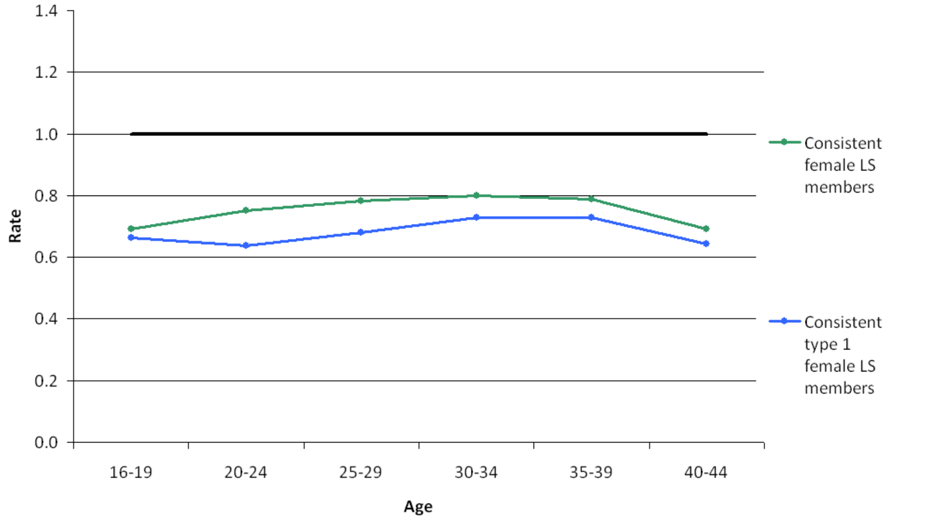
Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A53: Consistent cases - representation of LS births based on official statistics, 1999



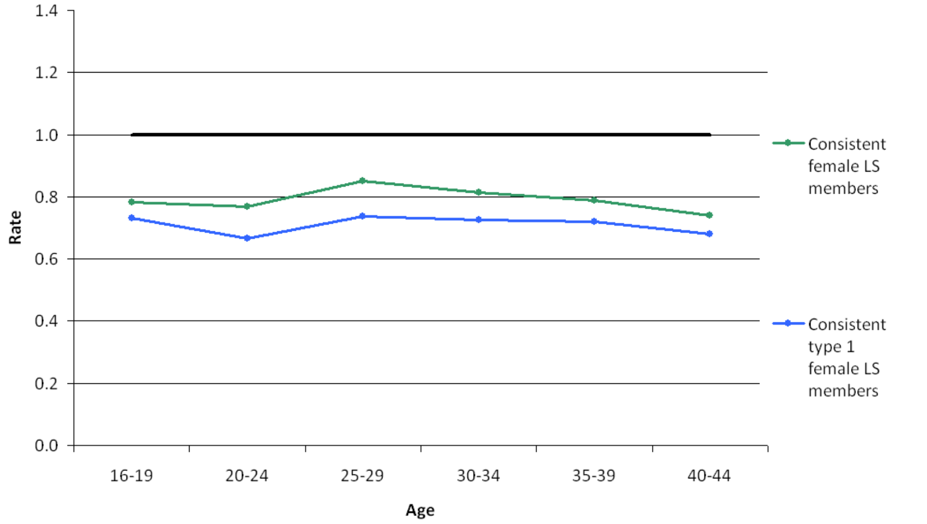
Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A54: Consistent cases - representation of LS births based on official statistics, 2000



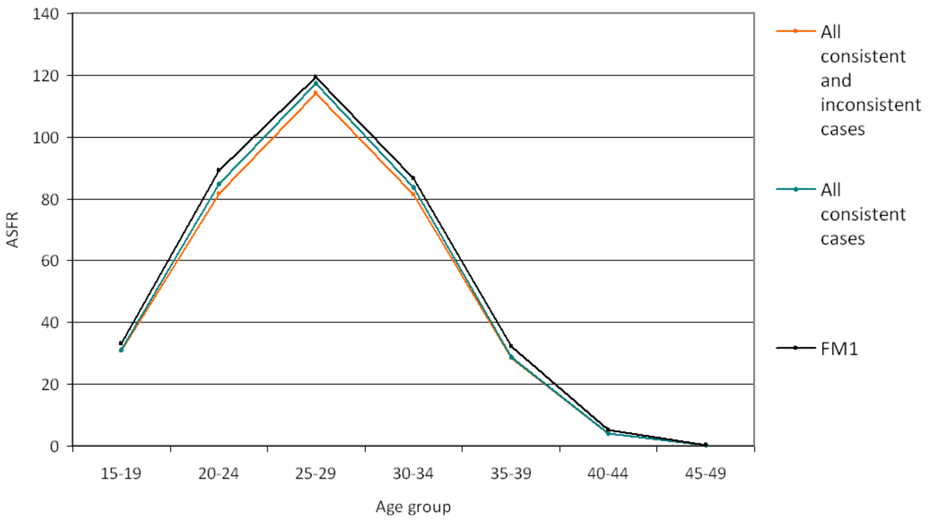
Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A55: Consistent cases - representation of LS births based on official statistics, 2001



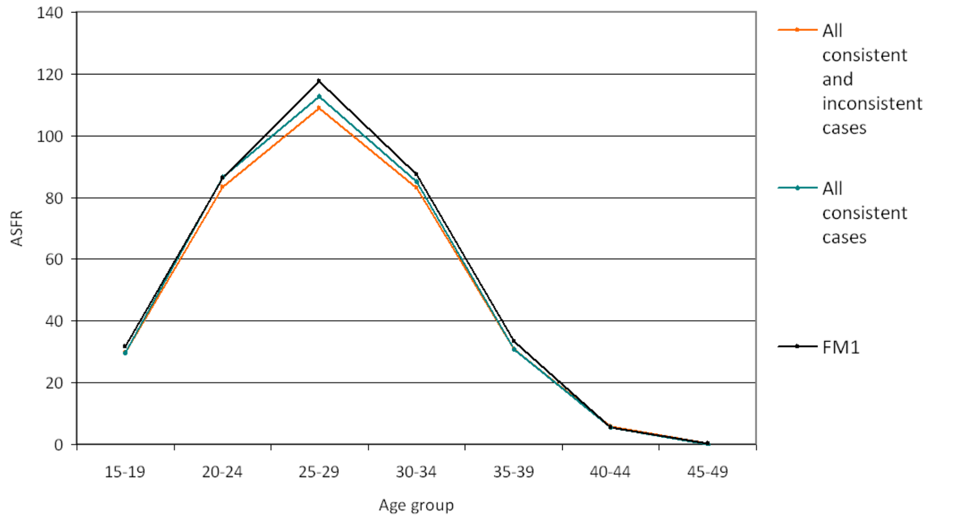
Own elaboration based on ONS LS, September 2010, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A56: Consistent and Inconsistent cases - Age-specific fertility rates, 1991



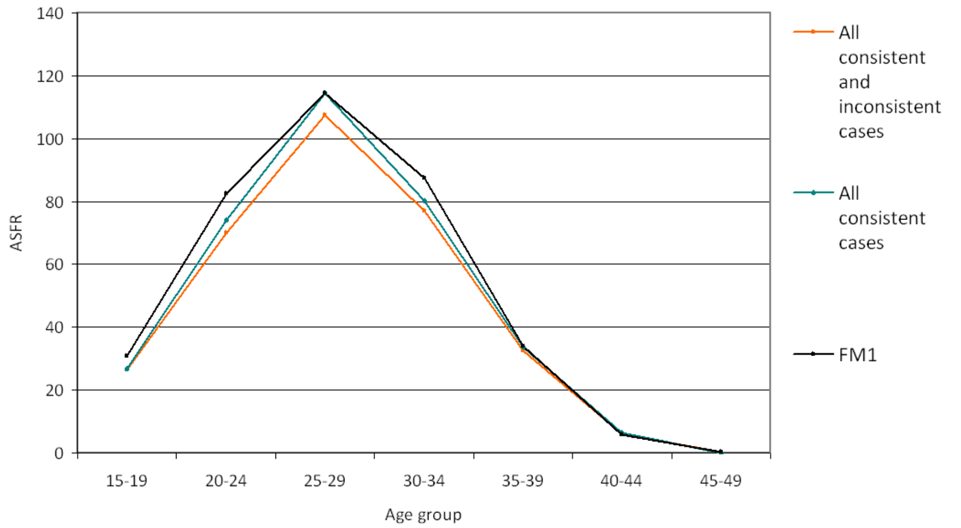
Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011.

Figure A57: Consistent and Inconsistent cases - Age-specific fertility rates, 1992



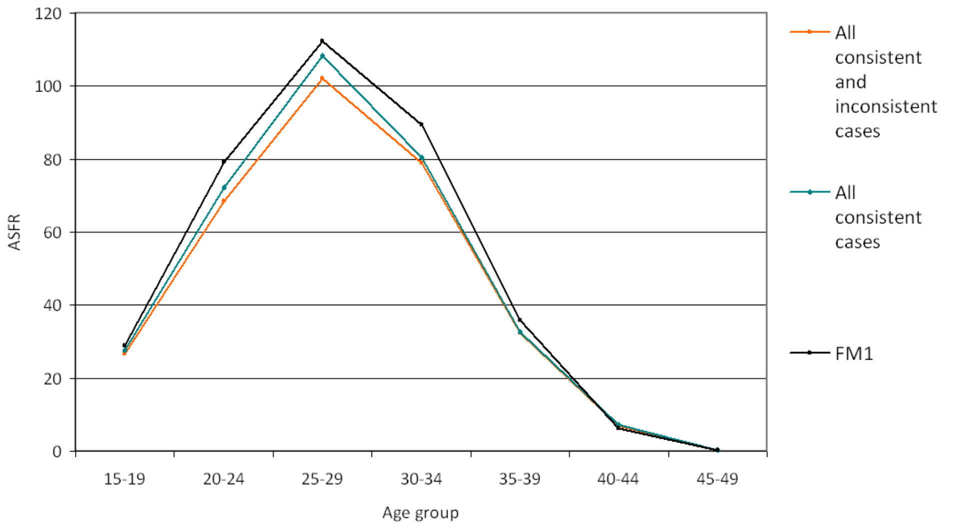
Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011.

Figure A58: Consistent and inconsistent cases - age-specific fertility rates, 1993



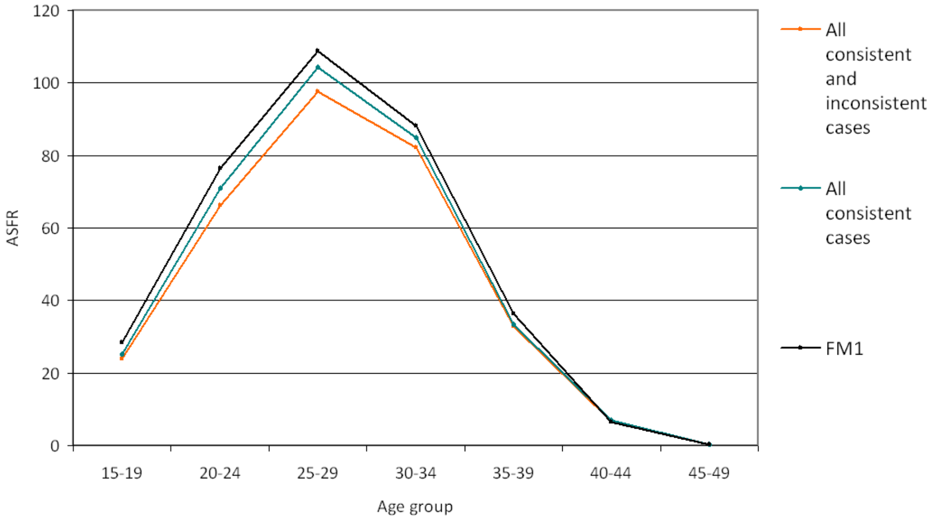
Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011.

Figure A59: Consistent and inconsistent cases - age-specific fertility rates, 1994



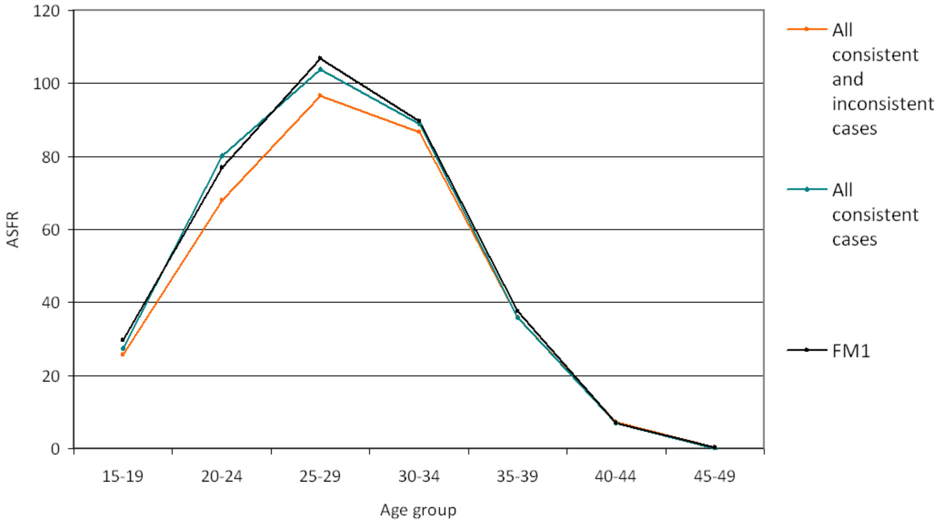
Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011.

Figure A60: Consistent and inconsistent cases - age-specific fertility rates, 1995



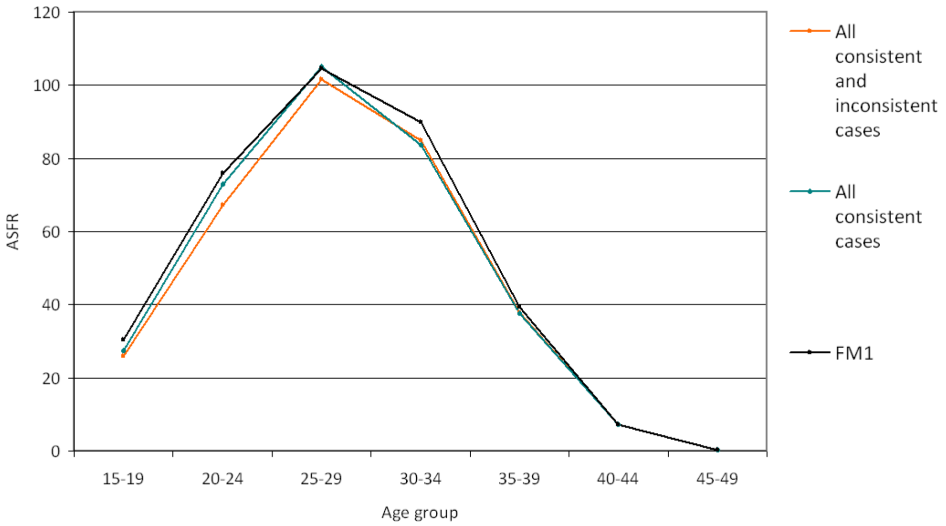
Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011.

Figure A61: Consistent and inconsistent cases - age-specific fertility rates, 1996



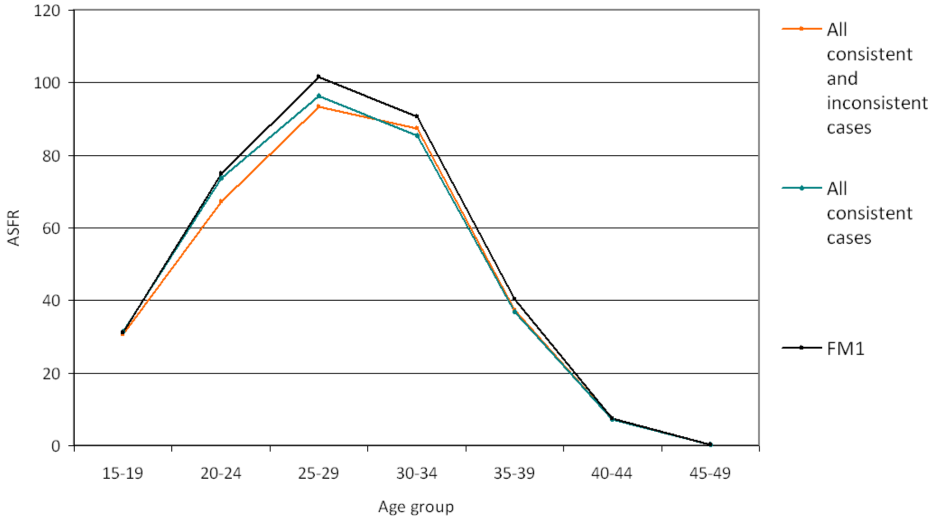
Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011.

Figure A62: Consistent and inconsistent cases - age-specific fertility rates, 1997



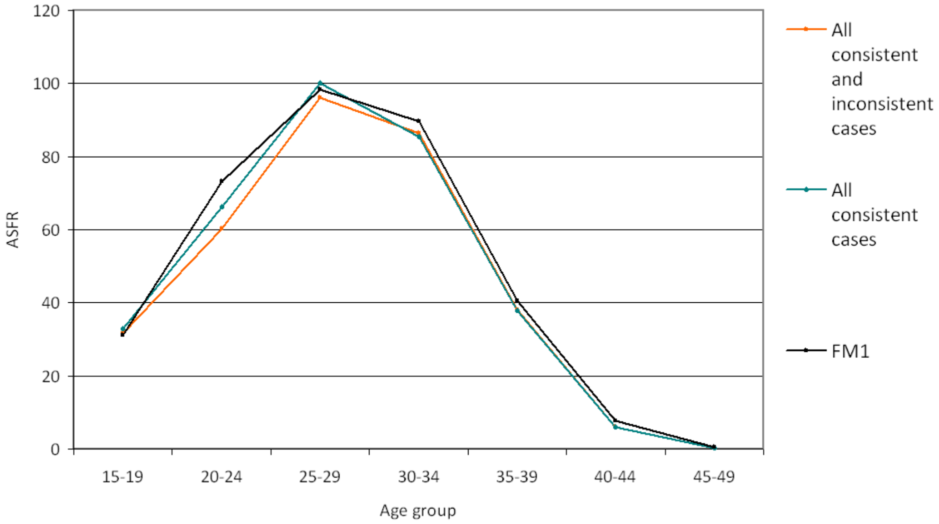
Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011.

Figure A63: Consistent and inconsistent cases - age-specific fertility rates, 1998



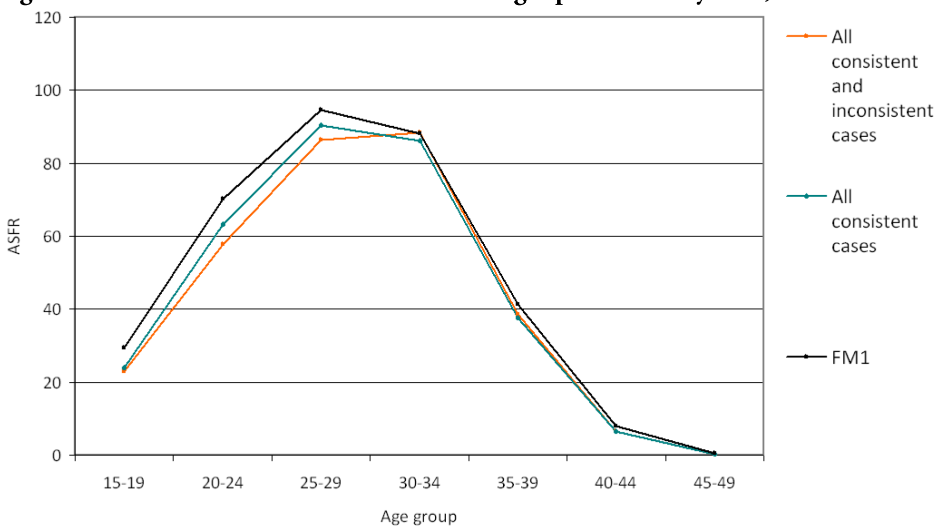
Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011.

Figure A64: Consistent and inconsistent cases - age-specific fertility rates, 1999



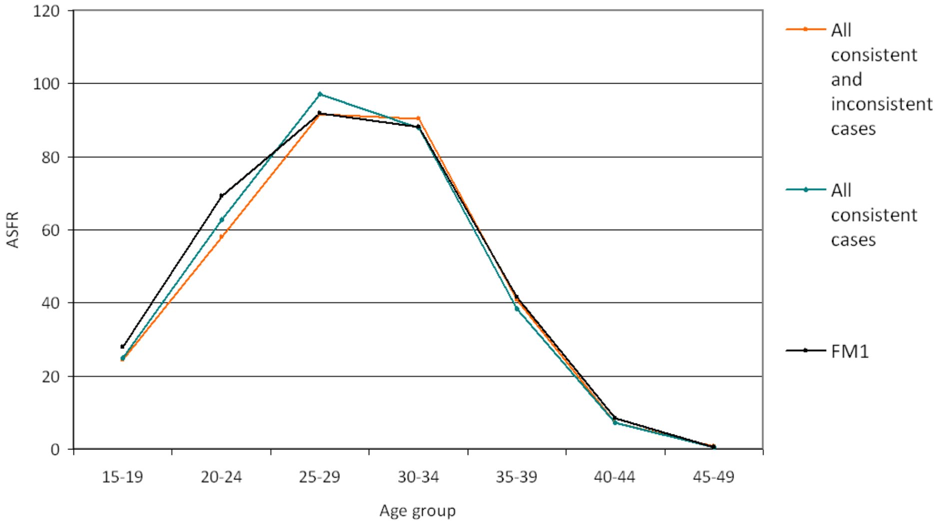
Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011.

Figure A65: Consistent and inconsistent cases - age-specific fertility rates, 2000



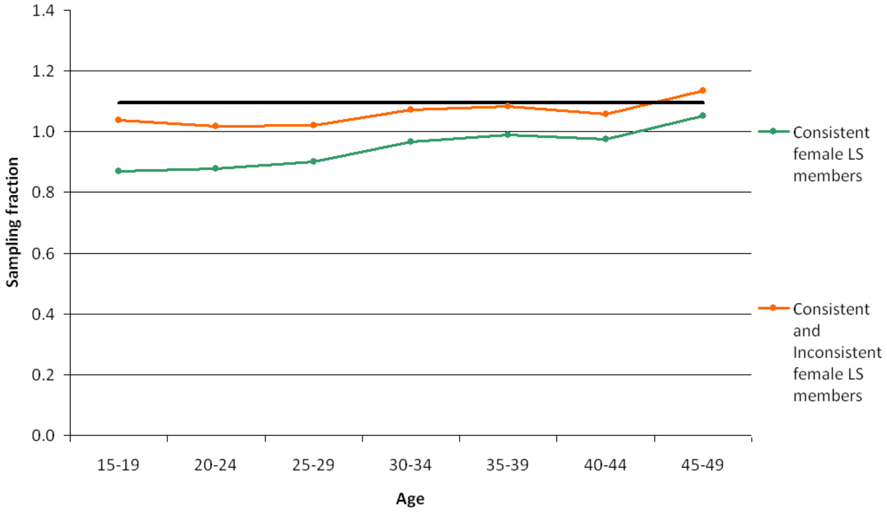
Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011.

Figure A66: Consistent and inconsistent cases - age-specific fertility rates, 2001



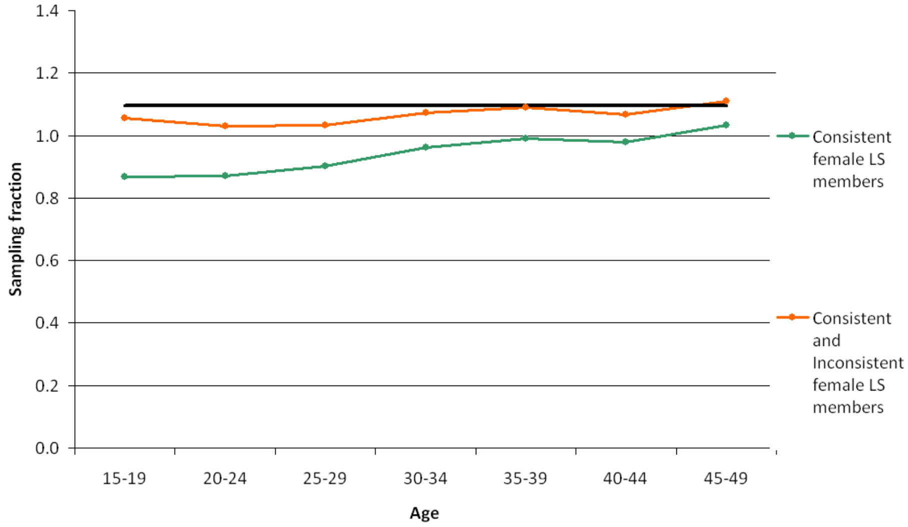
Own elaboration based on ONS LS, FM1 no.30, 2001, January 2011.

Figure A67: Consistent and inconsistent cases - sampling fraction of official statistics women by the LS, 1991



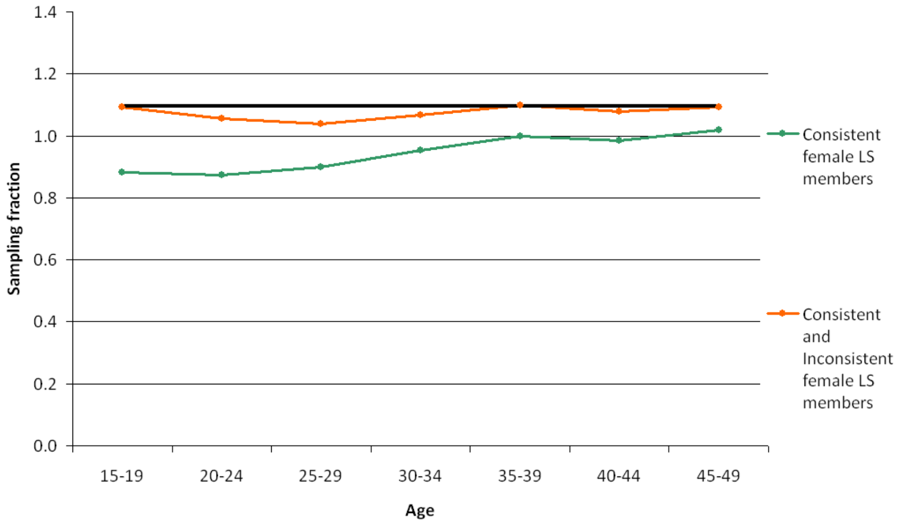
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A68: Consistent and inconsistent cases – sampling fraction of official statistics women by the LS, 1992



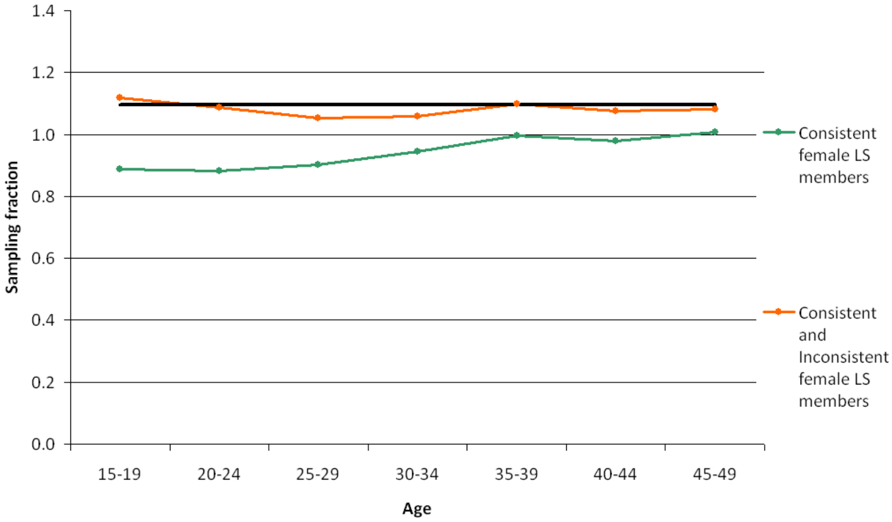
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A69: Consistent and inconsistent cases - sampling fraction of official statistics women by the LS, 1993



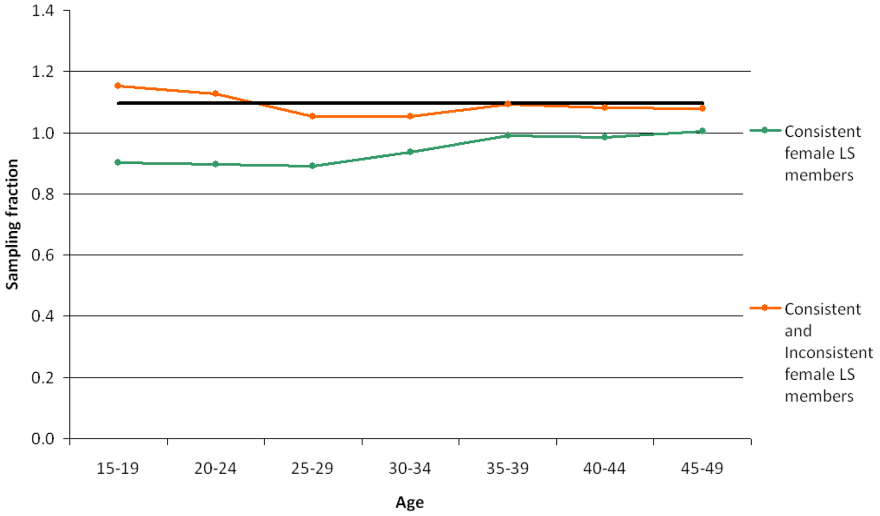
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A70: Consistent and inconsistent cases - sampling fraction of official statistics women by the LS, 1994



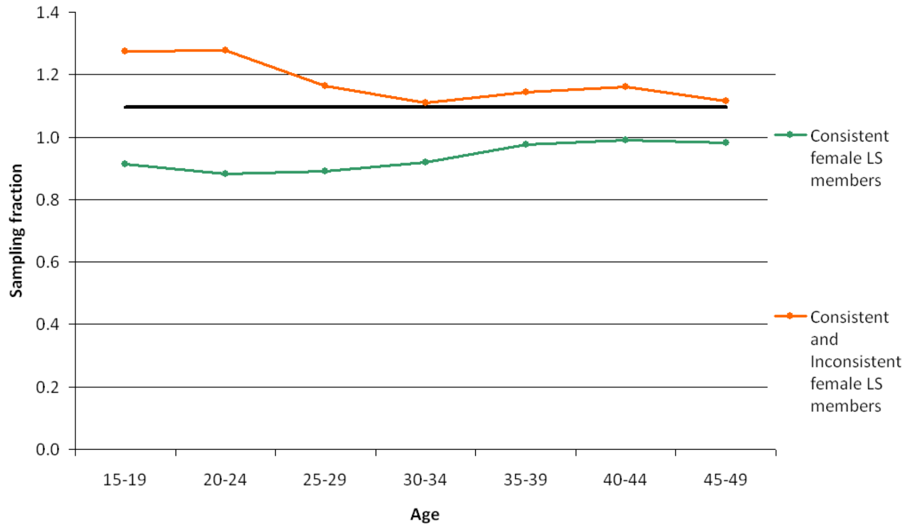
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A71: Consistent and inconsistent cases - sampling fraction of official statistics women by the LS, 1995



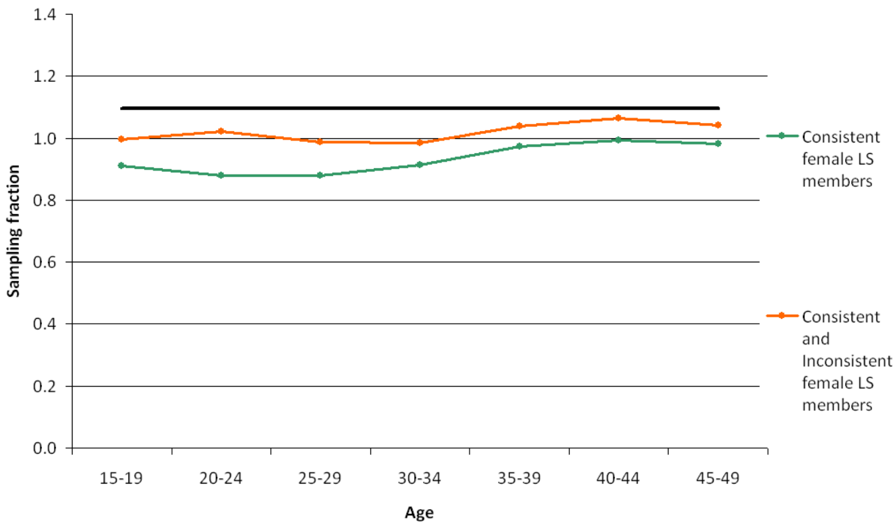
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A72: Consistent and inconsistent cases - sampling fraction of official statistics women by the LS, 1996



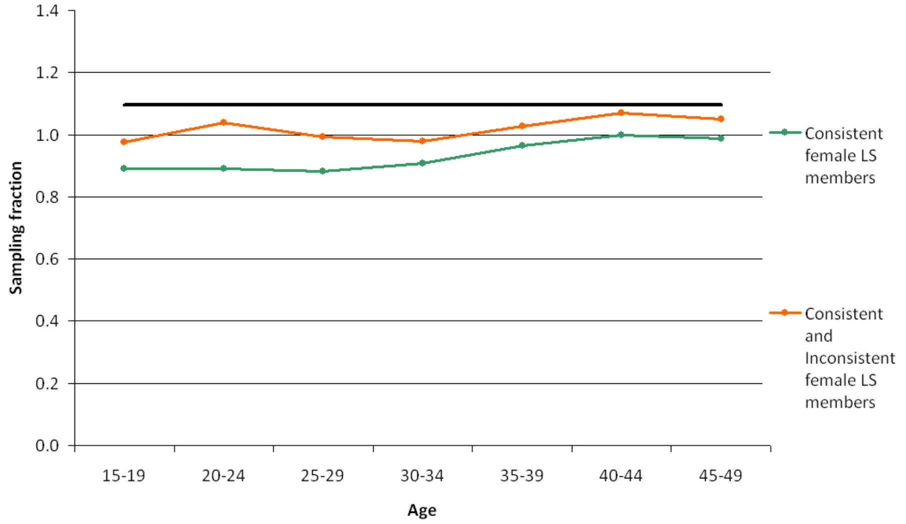
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A73: Consistent and inconsistent cases - sampling fraction of official statistics women by the LS, 1997



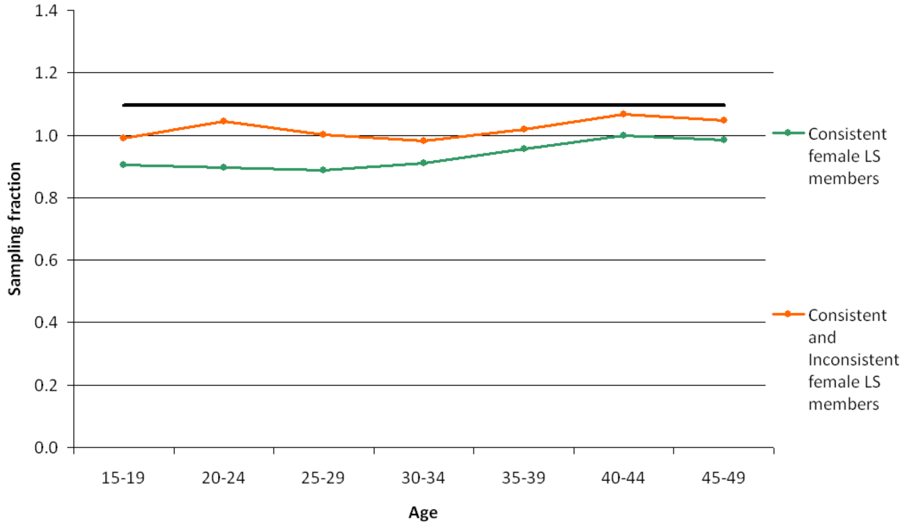
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A74: Consistent and inconsistent cases - sampling fraction of official statistics women by the LS, 1998



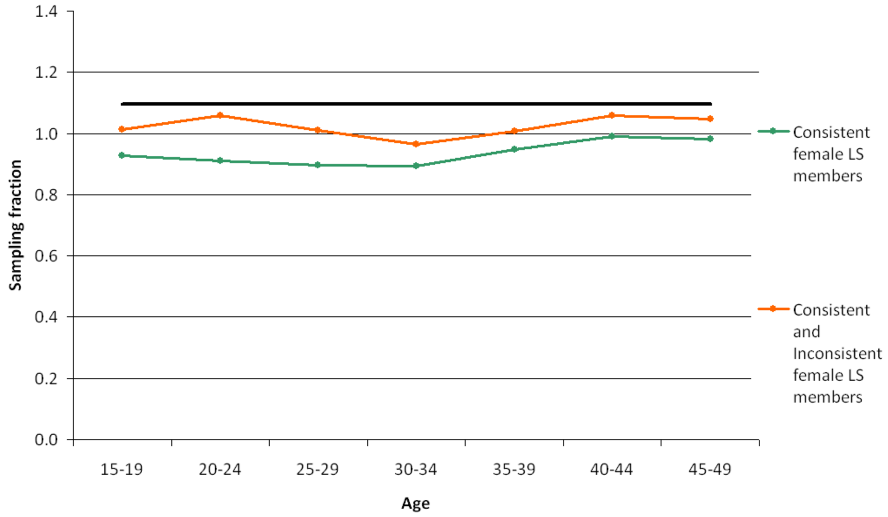
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A75: Consistent and inconsistent cases - sampling fraction of official statistics women by the LS, 1999



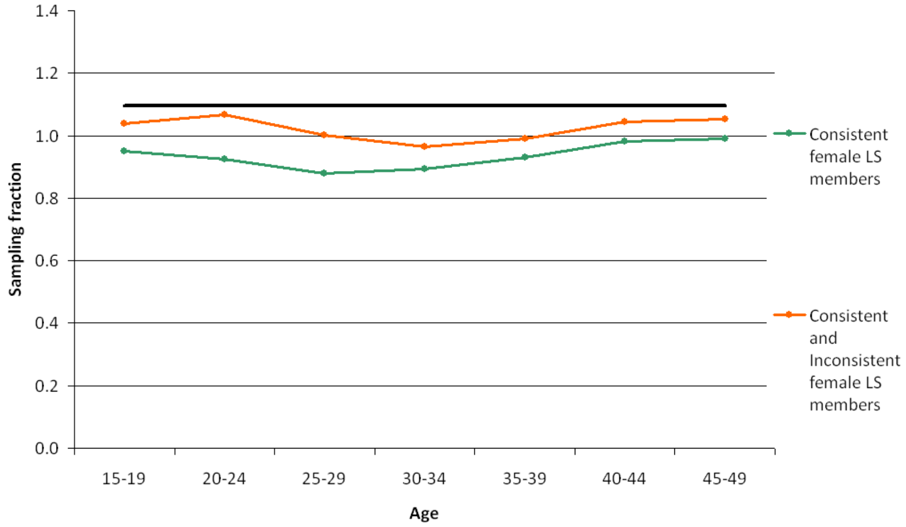
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A76: Consistent and inconsistent cases - sampling fraction of official statistics women by the LS, 2000



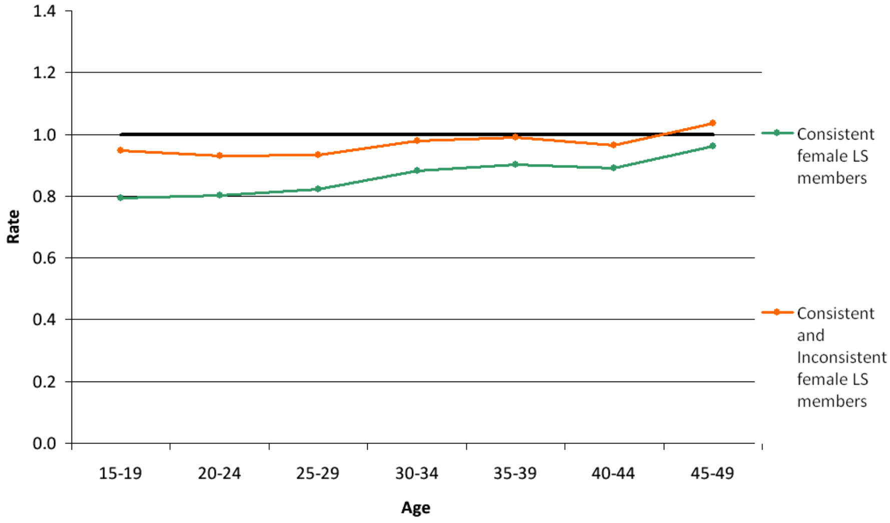
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A77: Consistent and inconsistent cases - sampling fraction of official statistics women by the LS, 2001



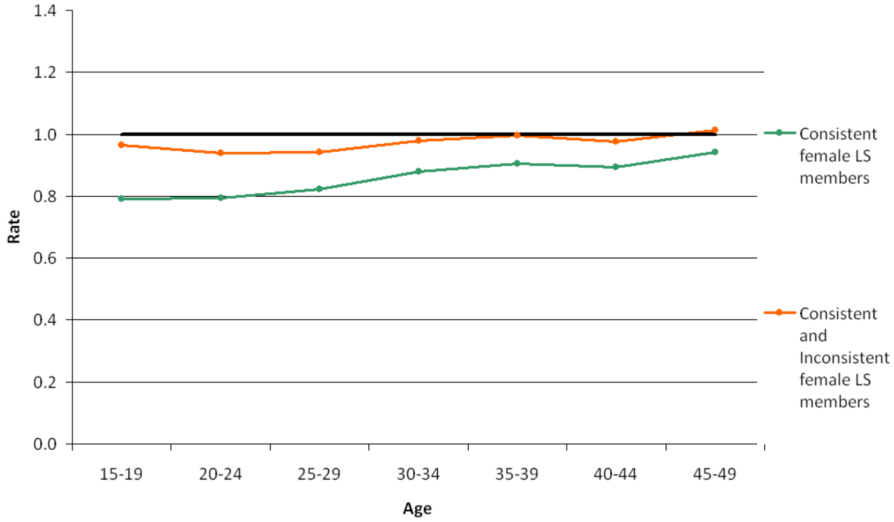
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A78: Consistent and inconsistent cases - representation of LS women based on official statistics, 1991



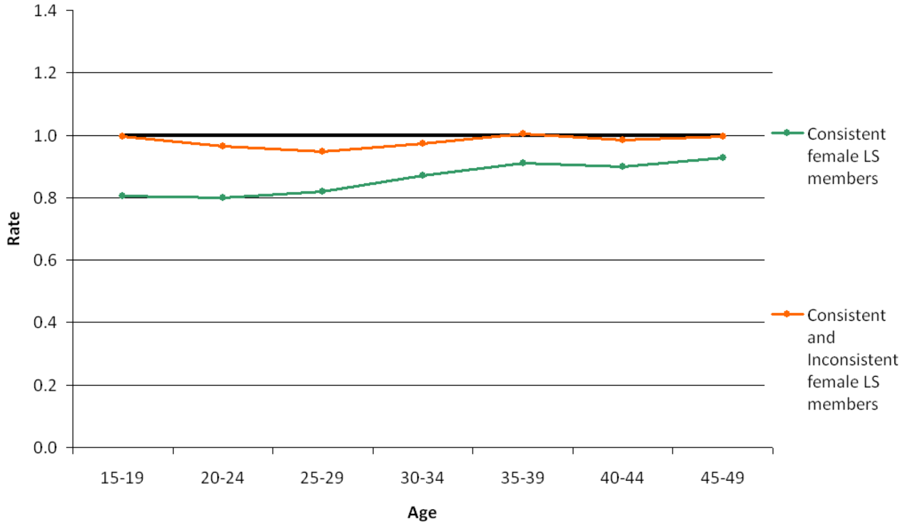
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A79: Consistent and inconsistent cases - representation of LS women based on official statistics, 1992



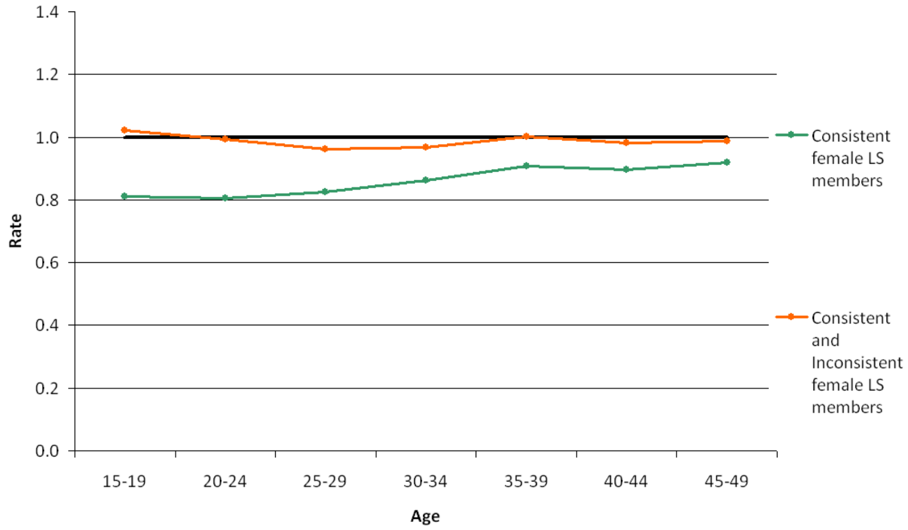
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A80: Consistent and inconsistent cases - representation of LS women based on official statistics, 1993



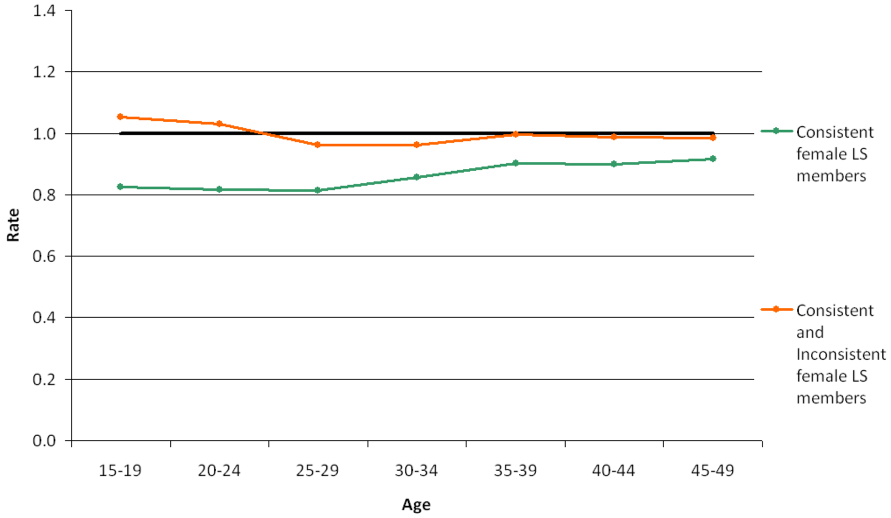
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A81: Consistent and inconsistent cases - representation of LS women based on official statistics, 1994



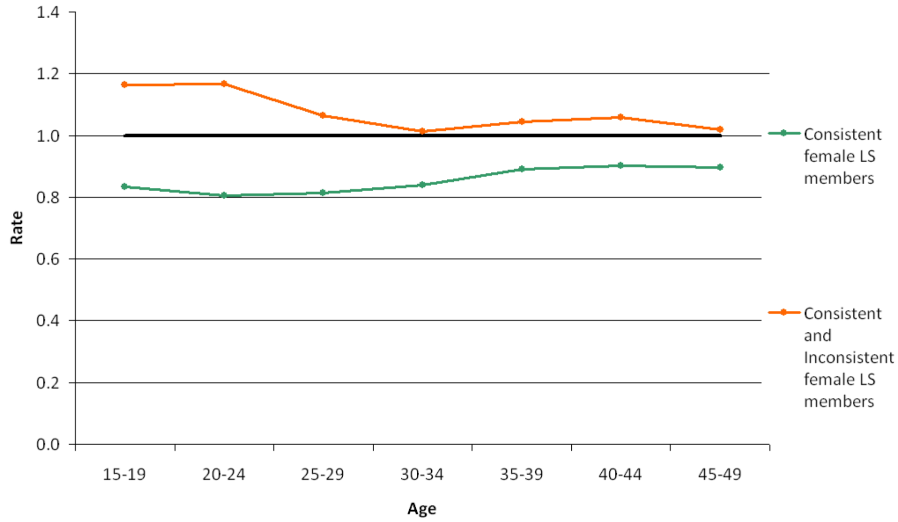
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A82: Consistent and inconsistent cases - representation of LS women based on official statistics, 1995



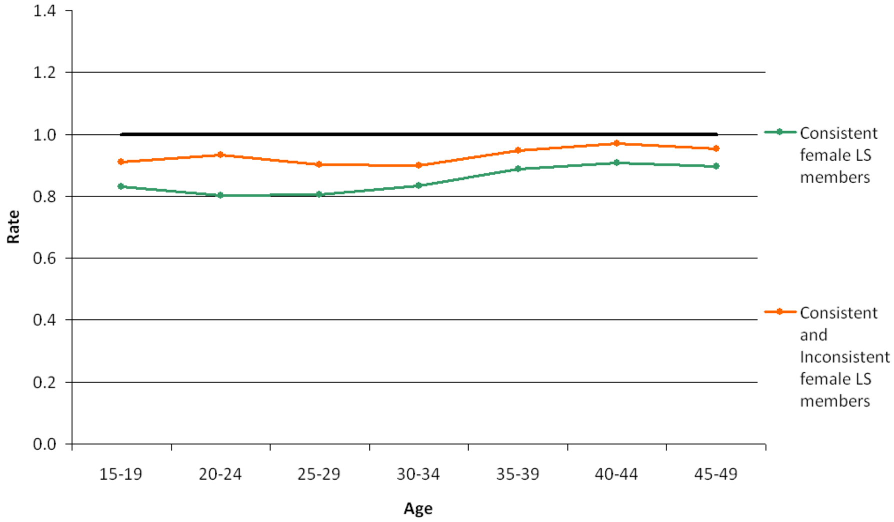
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A83: Consistent and inconsistent cases - representation of LS women based on official statistics, 1996



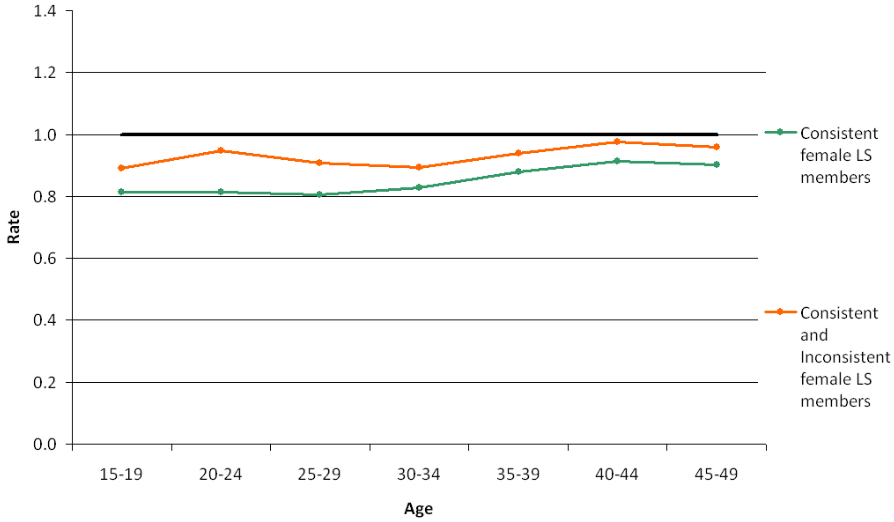
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A84: Consistent and inconsistent cases - representation of LS women based on official statistics, 1997



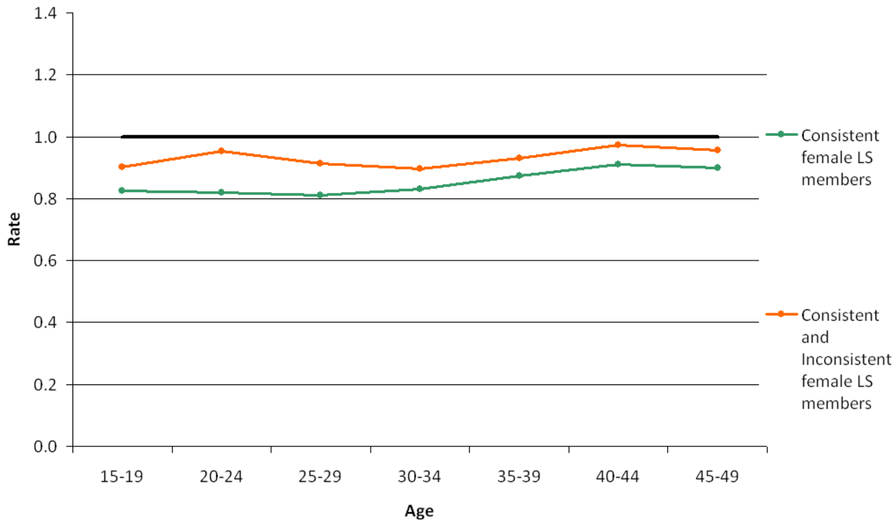
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A85: Consistent and inconsistent cases - representation of LS women based on official statistics, 1998



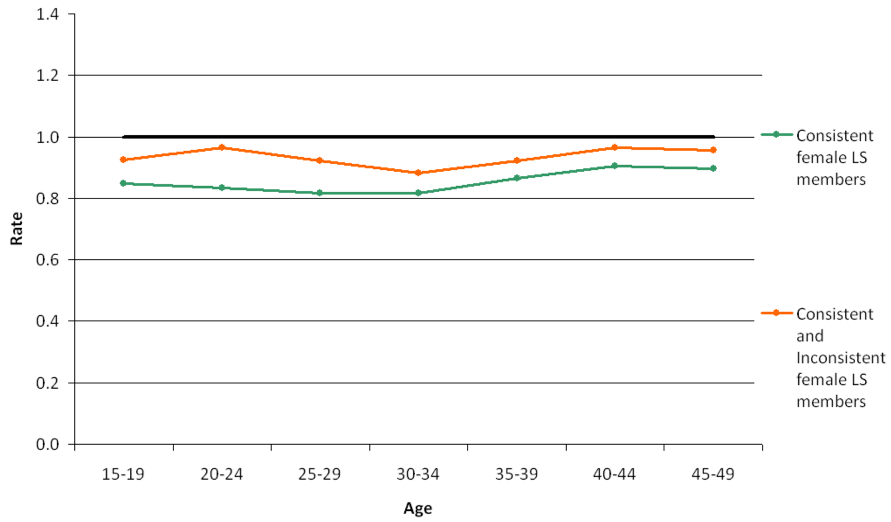
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A86: Consistent and inconsistent cases - representation of LS women based on official statistics, 1999



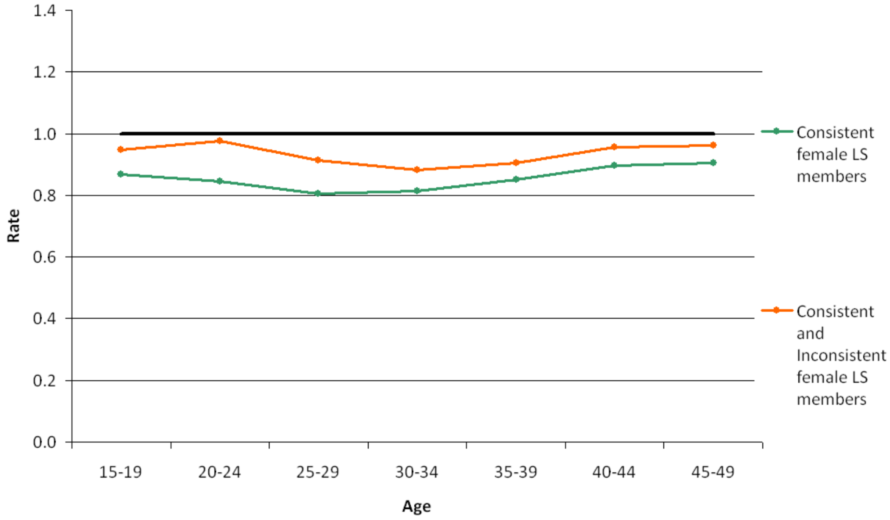
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A87: Consistent and inconsistent cases - representation of LS women based on official statistics, 2000



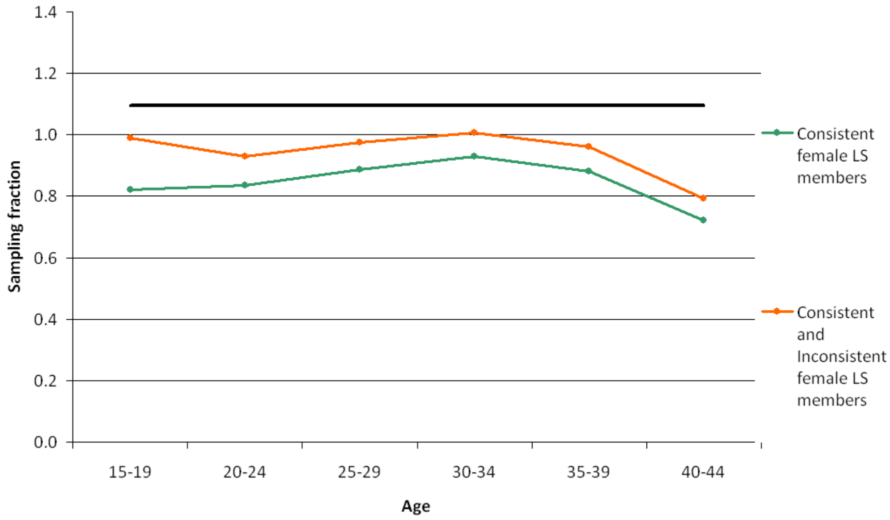
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A88: Consistent and inconsistent cases - representation of LS women based on official statistics, 2001



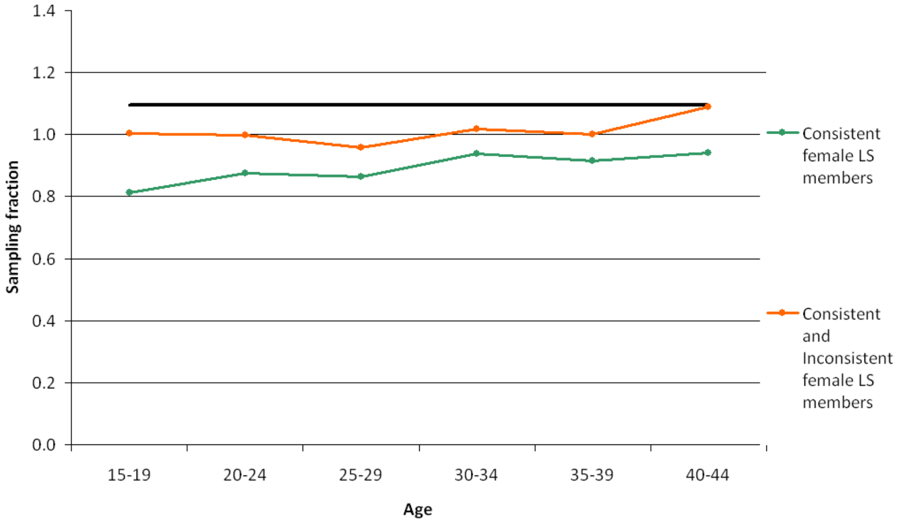
Own elaboration based on ONS LS, January 2011, ONS Mid-year estimates, Mid-1991 to Mid-2000 LA Population Studies: 07/10/04, June 2010.

Figure A89: Consistent and inconsistent cases - sampling fraction of official statistics births by the LS, 1991



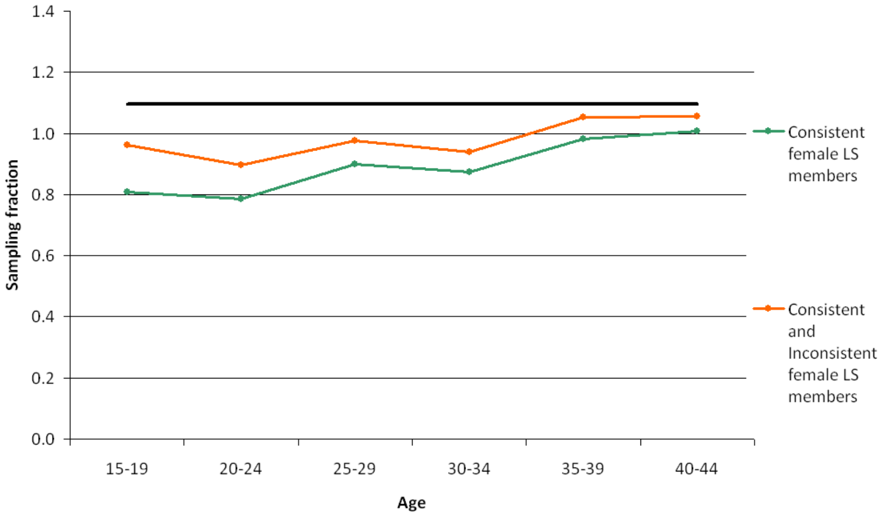
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A90: Consistent and inconsistent cases - sampling fraction of official statistics births by the LS, 1992



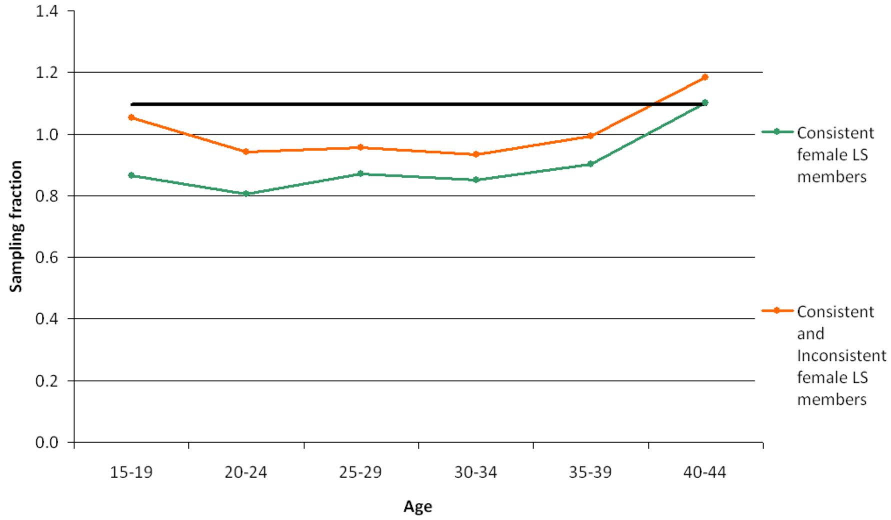
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A91: Consistent and inconsistent cases - sampling fraction of official statistics births by the LS, 1993



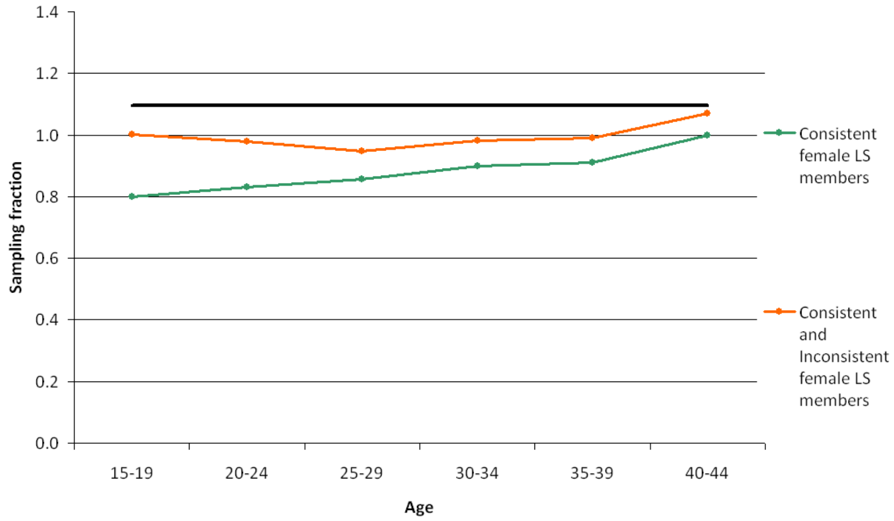
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A92: Consistent and inconsistent cases - sampling fraction of official statistics births by the LS, 1994



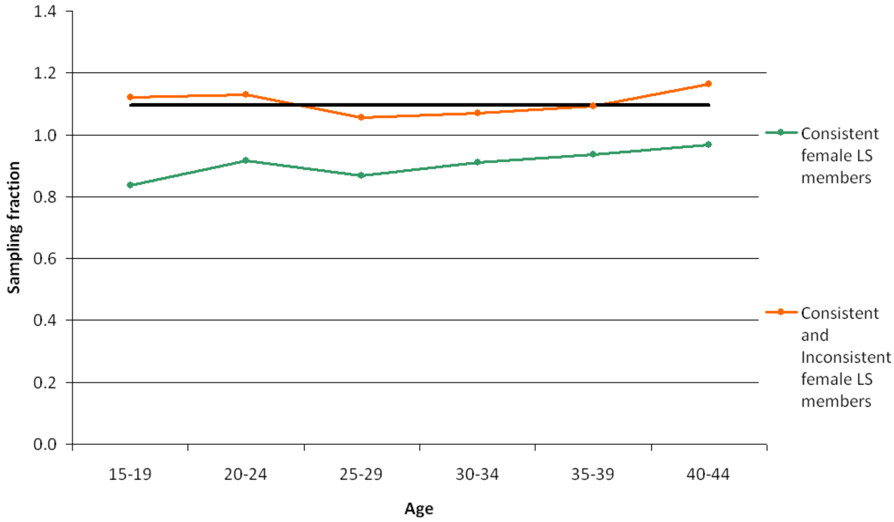
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A93: Consistent and inconsistent cases - sampling fraction of official statistics births by the LS, 1995



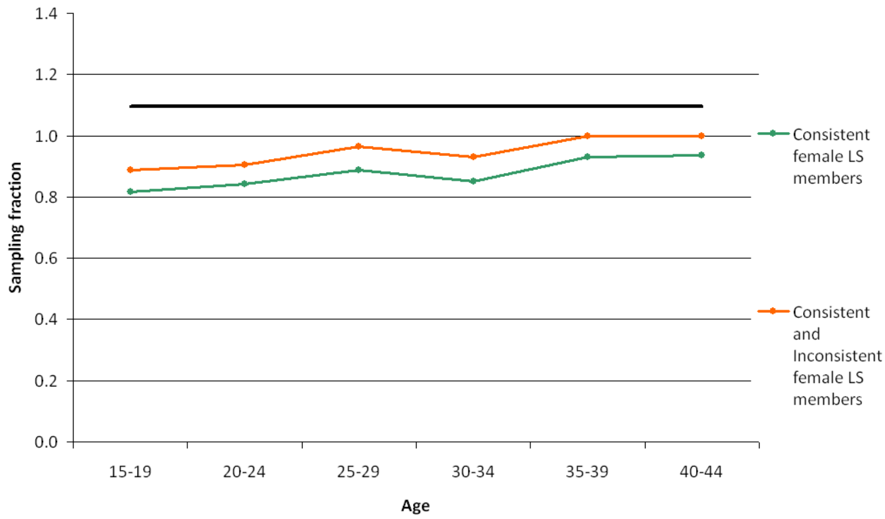
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A94: Consistent and inconsistent cases - sampling fraction of official statistics births by the LS, 1996



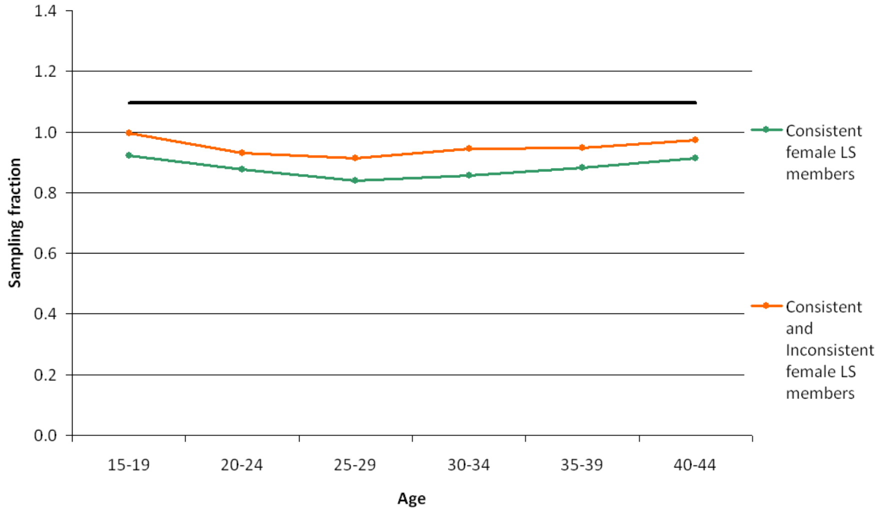
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A95: Consistent and inconsistent cases - sampling fraction of official statistics births by the LS, 1997



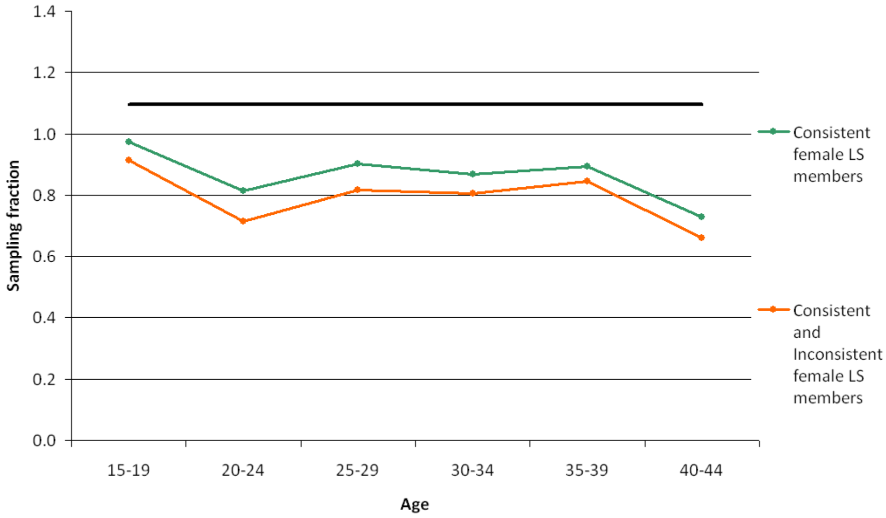
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A96: Consistent and inconsistent cases - sampling fraction of official statistics births by the LS, 1998



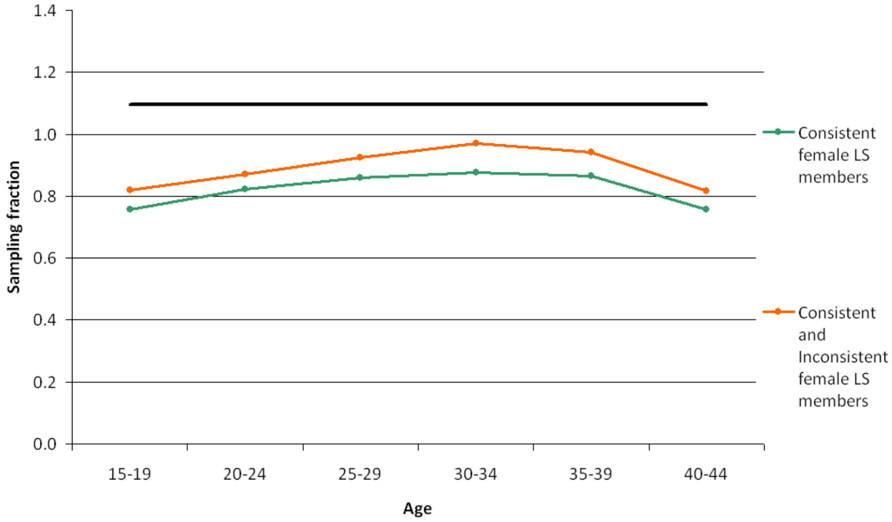
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A97: Consistent and Inconsistent cases - sampling fraction of official statistics births by the LS, 1999



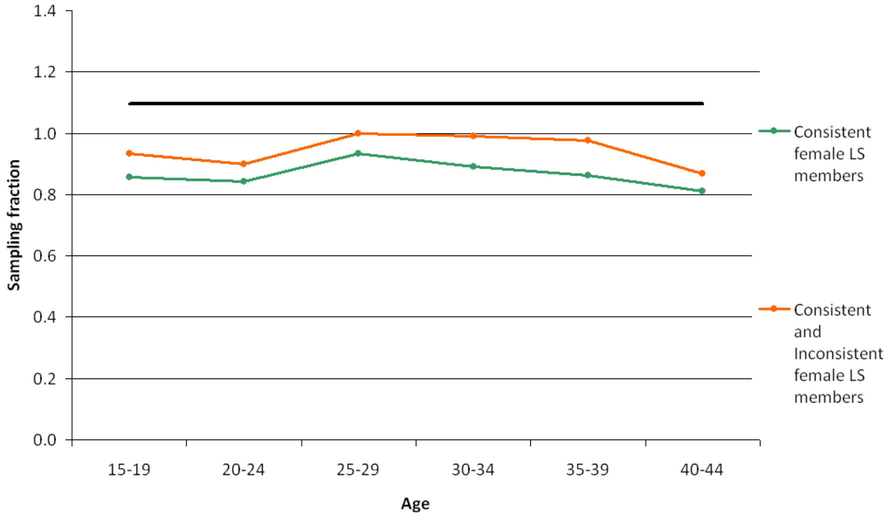
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A98: Consistent and inconsistent cases - sampling fraction of official statistics births by the LS, 2000



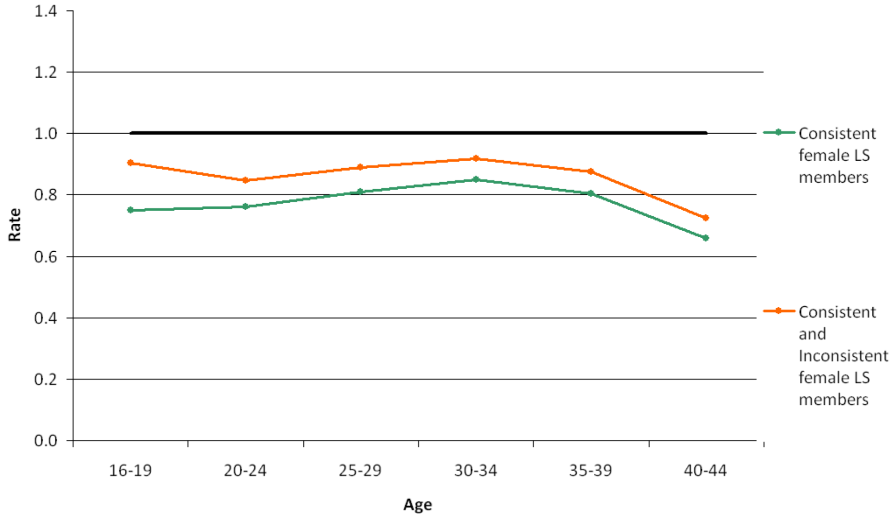
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A99: Consistent and inconsistent cases - sampling fraction of official statistics births by the LS, 2001



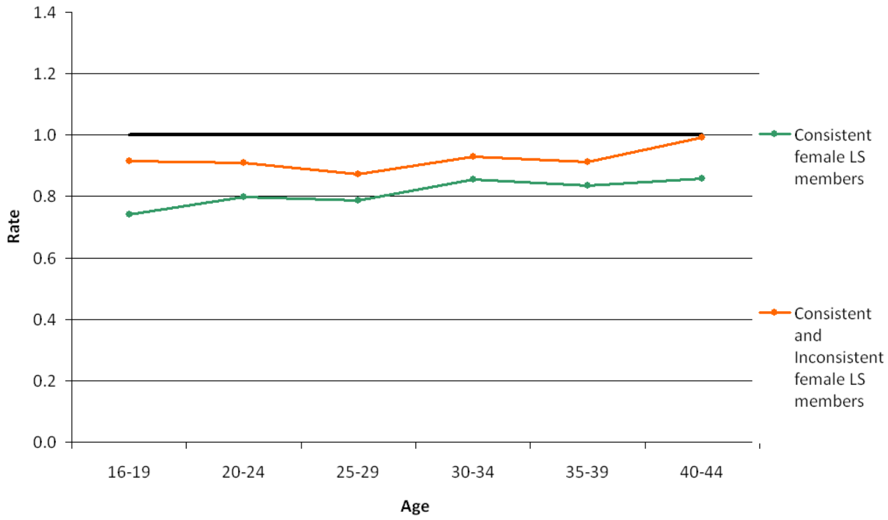
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A100: Consistent and inconsistent cases - representation of LS births based on official statistics, 1991



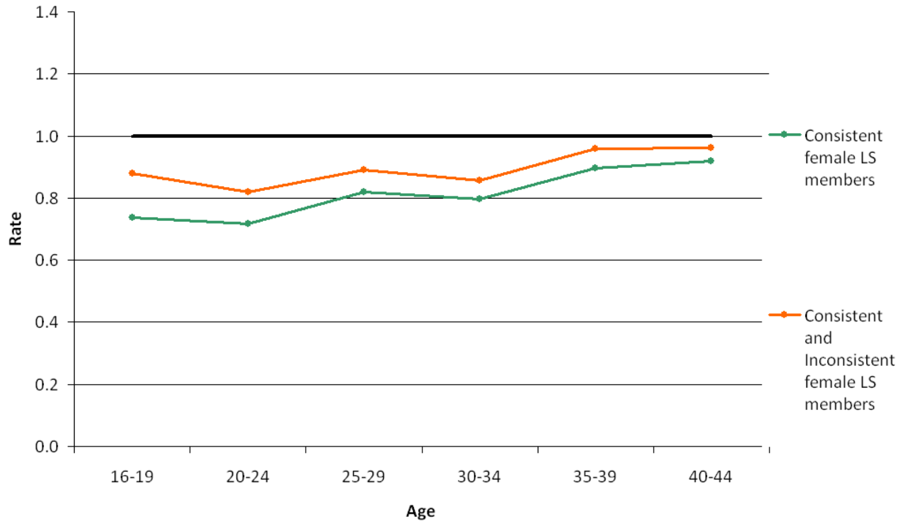
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A101: Consistent and inconsistent cases - representation of LS births based on official statistics, 1992



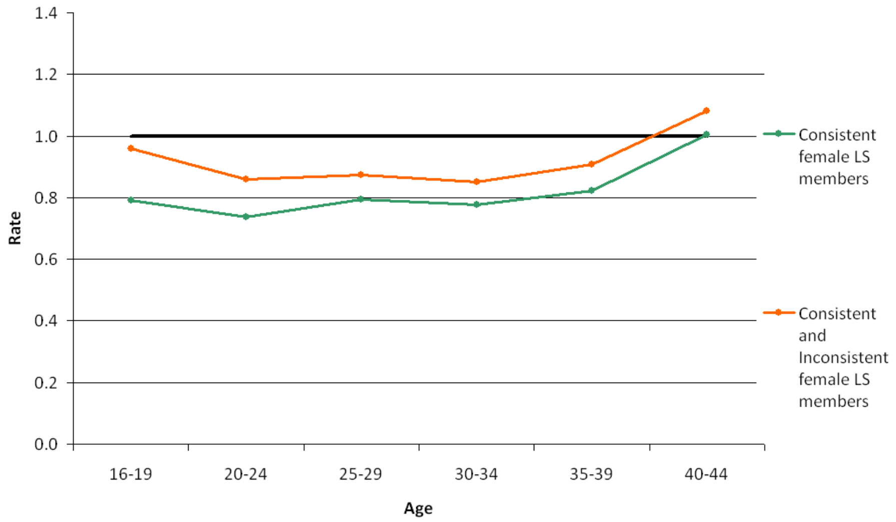
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A102: Consistent and inconsistent cases - representation of LS births based on official statistics, 1993



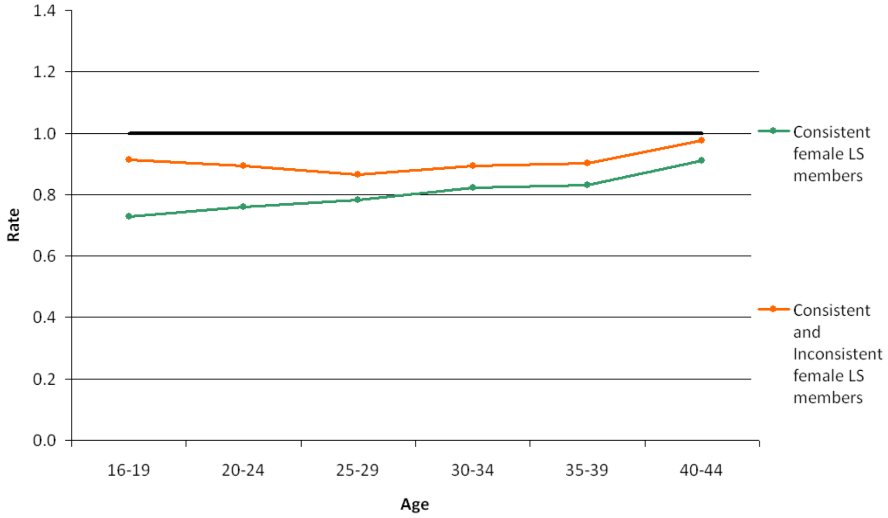
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A103: Consistent and inconsistent cases - representation of LS births based on official statistics, 1994



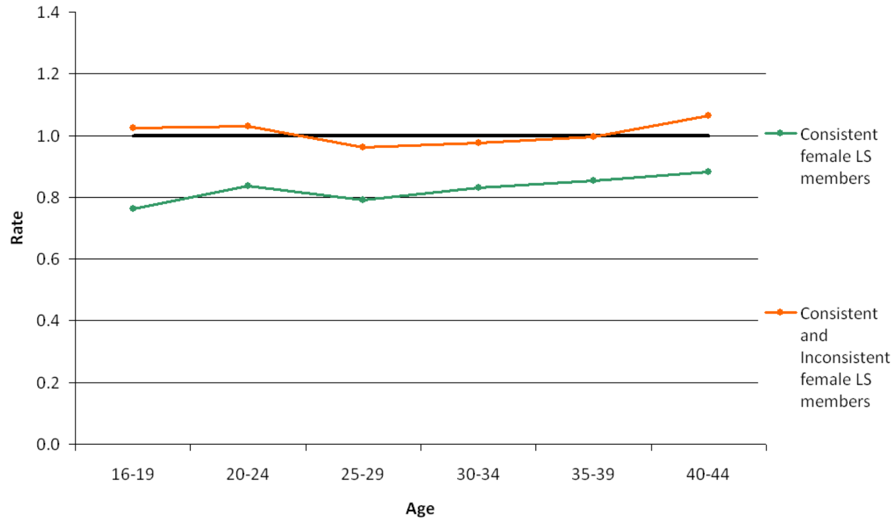
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A104: Consistent and inconsistent cases - representation of LS births based on official statistics, 1995



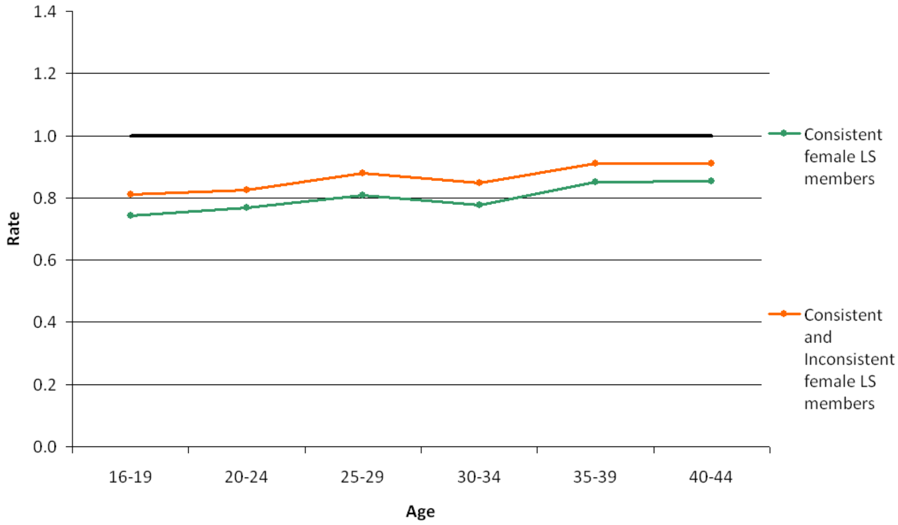
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A105: Consistent and inconsistent cases - representation of LS births based on official statistics, 1996



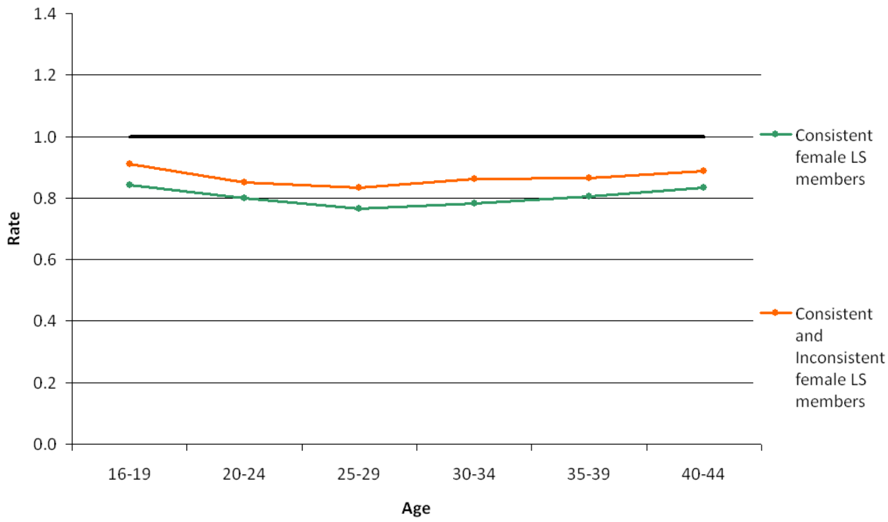
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A106: Consistent and inconsistent cases - representation of LS births based on official statistics, 1997



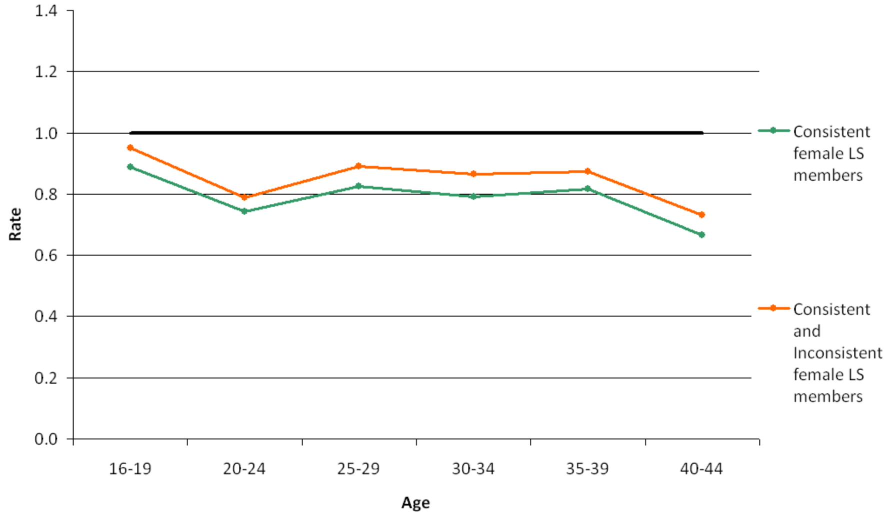
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A107: Consistent and inconsistent cases - representation of LS births based on official statistics, 1998



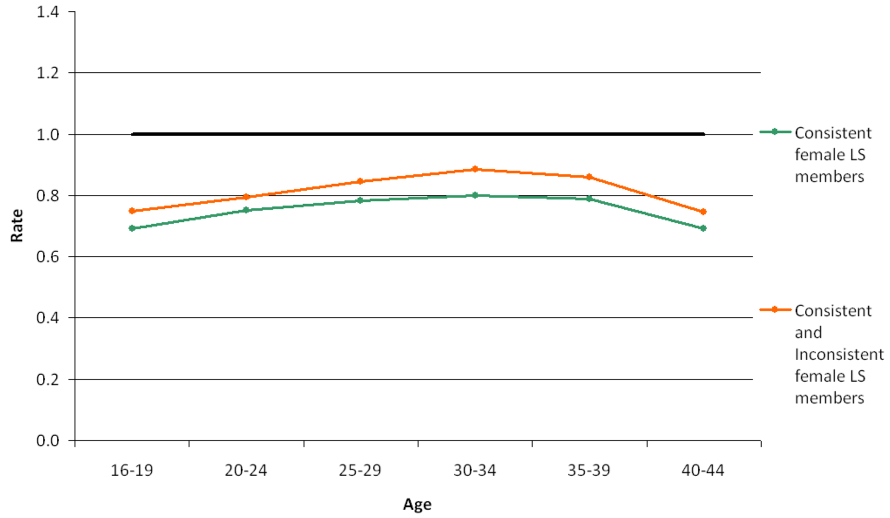
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A108: Consistent and inconsistent cases - representation of LS births based on official statistics, 1999



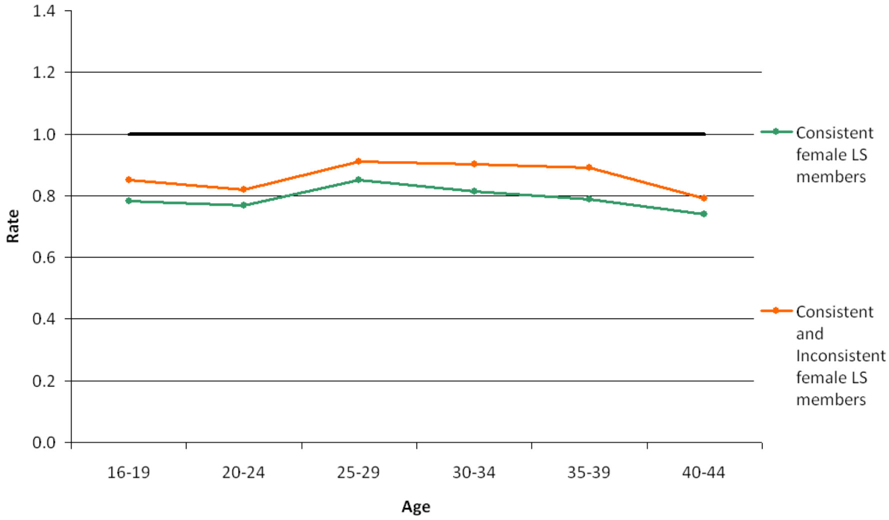
Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A109: Consistent and inconsistent cases - representation of LS births based on official statistics, 2000



Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.

Figure A110: Consistent and inconsistent cases - representation of LS births based on official statistics, 2001



Own elaboration based on ONS LS, January 2011, ONS FM1 volume - Births: 1938-2004, Maternities, Age of mother, a. all maternities, June 2010.