

Sex ratio patterns in population estimates

Steve Smallwood and Sofie De Broe
Office for National Statistics

Introduction

This paper presents research on the national sex ratio pattern observed in England and Wales (EW) in population estimates following the 2001 Census. The sex ratio is defined as the number of males per 100 females. Sex ratios are determined by the three fundamental factors of demographic change: births, deaths, and migration, though the respective influences of these drivers vary by age. Sex ratios for children are primarily determined by the sex ratio at birth; for a large western developed country this is stable at around 105 boys per 100 girls.¹ National sex ratios at the younger working ages are primarily determined by international migration. At older ages (from around age 55 onwards) the longer life expectation of women compared with men drives the sex ratio. The sex ratio is independent of the absolute numbers of males and females in a large population and is therefore often used as a quality measure of data by age and sex.²

Since 2001 the sex ratio in the mid-year estimates (MYEs) has produced a pattern which has raised questions in terms of its plausibility. This paper presents the research on the sex ratio pattern observed in EW over recent censuses and for population estimates over the current intercensal period. Its aim is to understand the drivers of this sex ratio and provide evidence on its plausibility. The discussion addresses remaining questions on recent sex ratio patterns and suggests ways forward in order to prepare for the 2011 Census and beyond.

The 2001 Census was the first Census that was adjusted for underenumeration, enabled by a large coverage survey.³ After the Census, analysis suggested some limitations in a few areas where it was not able sufficiently to adjust for exceptional circumstances. Additions to the population were derived from the ONS Longitudinal Study (LS) and

The mid-year estimates (MYEs) resulting from the 2001 Census have a sex ratio pattern differing from patterns in previous estimates. Research has been undertaken to explain the observed sex ratio pattern in England and Wales (EW), covering recent censuses and population estimates during the current intercensal period. The aims of this paper are to present some of the research explaining this sex ratio pattern and, secondly, suggest ways to prepare for the 2011 Census and beyond. A number of approaches have been taken to understand the drivers of the sex ratio pattern and provide evidence on plausibility. The Patient Register Data (PRD) was examined as a potential comparator. Sex ratios in other countries were compared with the EW sex ratio patterns. An accounting exercise for different explanatory scenarios around the remaining unexplained difference between estimates and 2001 Census was undertaken. Finally, demographic analyses examine trends in sex ratios among the migrant stock of the EW and foreign born populations to put into context the sex ratios in MYEs following the 2001 Census.



local authority studies⁴ and applied to the mid-2001 population estimates. In the adjusted 2001 population estimates, there is a very sharp drop in the sex ratio after age 18 from around 105 men per 100 women to around 100 men per 100 women. An additional feature of the EW sex ratio pattern in the population estimates is that the sharp drop observed around age 18 in 2001 MYEs appears at one year older in each successive year's MYEs (the pattern ages forward), as shown in **Figure 1**.

This paper is structured as follows. After an initial discussion of current sex ratios in mid-year estimates, sex ratios in a different administrative population dataset, the Patient Register Data (PRD), are examined and a potential explanation provided for the patterns found there. Evidence of sex ratios in other countries, including the other UK countries, is considered, and then an accounting exercise is undertaken presenting two explanatory scenarios. Finally sex ratios of migrant stocks by country of birth are presented. This analysis looks at the sex ratio of the non-EW born found in the Census and, given EW births and mortality rates, the implied sex ratio of the EW born abroad. This helps us to understand whether the sex ratio is being driven by sex imbalances in net immigration from overseas or net emigration of EW born. Work in progress and ways forward towards 2011 and beyond are then discussed.

The paper concentrates mainly on the sex ratios found in MYEs and PRD. However, ratios result from the underlying numbers of people by age and sex and this information is presented for reference in Appendix A.

The 2001 Census and mid-year population estimates

The mid-2001 population estimates based on the 2001 Census for the population of EW initially estimated the population at 1.1 million lower than the MYE rolled forward from the 1991 Census. Over half of the difference was explained by two elements: that additions to the population as a result of under enumeration in the 1991 Census were too high; and, errors in the estimation of migration to and from the UK during the 1990s.⁵ Further analysis of the ONS Longitudinal Study (ONS LS), exercises in two local authorities (LAs) (Manchester and Westminster⁶), and further studies of other LAs,⁷ identified a total of 275,000 people that were added to the mid-2001 population estimates partly to account for definitional differences between the 2001 Census and the mid-year estimates.^{8,9} There remains an unexplained difference of 209,000, which is 0.4 per cent of the total EW population. A later section of the paper explores this unexplained difference in more depth.

Population estimates for EW made prior to the 2001 Census suggested that men continued to outnumber women until their late forties, when the sex ratio started to fall. Estimates based on the 2001 Census show women outnumbering men as early as age 22. This trend had been observed in the 1991 Census but there was little belief in this finding until it was found again in the 2001 Census after the adjustments for under enumeration. A previous article in this journal has described the sex ratio patterns seen across recent censuses¹⁰, and points out that the dipping trend in the 2001 sex ratios is consistent with the trends observed in the 1981 and 1991 Censuses.¹⁰ The sex ratio for the 2001 MYEs (based on the 2001 Census and containing additions made in 2001 and 2003) shows a clear dip from age 18 onwards as shown by the thick black line in Figure 1. Additionally, the population estimates for subsequent years show that the sharp drop in the sex ratio is ageing forward.

The dip in the sex ratio at ages 18–19

The 2001 sex ratio pattern could be explained in a number of ways. A dip that ages on could only be explained by a particular sex imbalance in migration around ages 18–19 as a 'one-off' event related to the year 2000–2001. No evidence of this has been found in migration data. If the sharp dip in the sex ratio around age 18 is a 'real' effect of young men migrating around the age of 18–19 (or greater numbers of females than males arriving from abroad), then one might expect some male migrants returning at a later age (or female migrants departing), but, the dip to remain at around age 19 and be reproduced in subsequent MYEs at the same age.

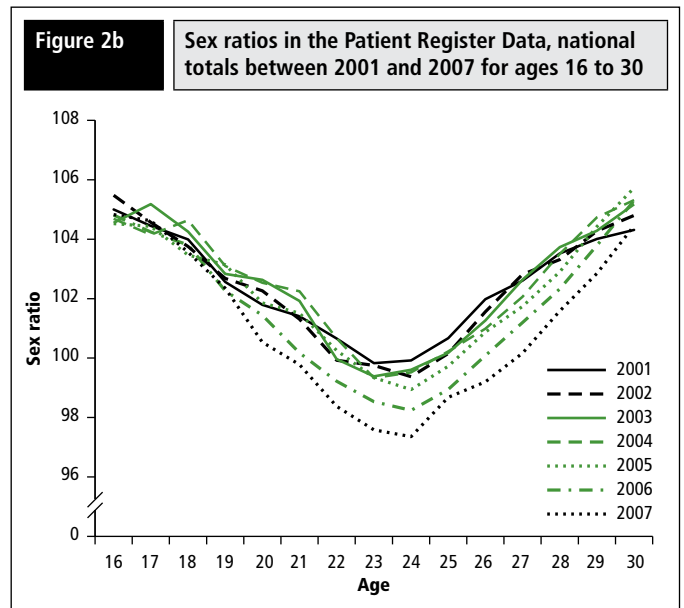
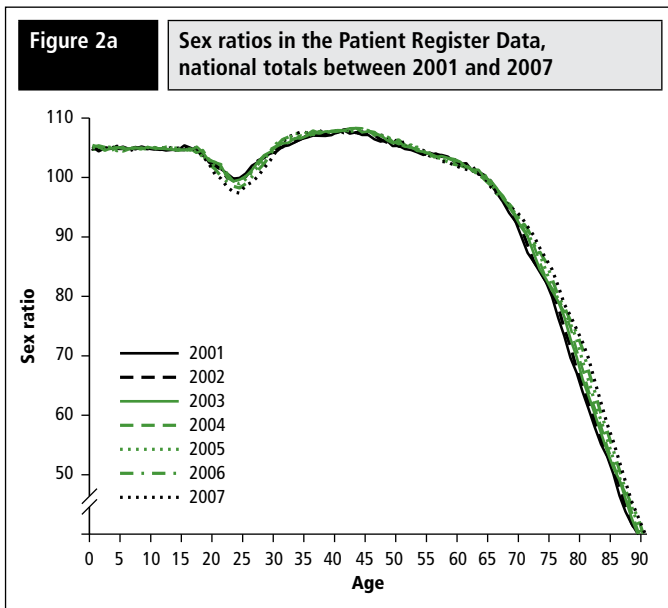
Another possible explanation for the cliff edge, and its moving forward pattern, could relate to population definitions, whereby young men are underrepresented in the population estimates for 2001 owing to the absence of a technique for adding returning short-term emigrants, not enumerated in the census. The pattern could also be partly explained if the widely recognised problem of under-enumeration of young males had not been fully addressed through census adjustments. A lot of work has been undertaken to address the unexplained difference; however, there is still a gender imbalance in the remaining 209,000. The explanation for the sharp dip in the sex ratio in the 2001 MYEs may be a combination of all these effects. While the dip may be a combination of causes, the ageing on is an indication that there is an issue in the methods or data sources used for population estimates. This indicates the need to understand further: migrant sex ratios; definitional issues in the Census; and, the residual issues in the unaccounted for 209,000 difference between rolled forward MYEs and Census. The remainder of this paper describes some of the analyses undertaken so far. There is more work being carried out than described here and this, together with any possible way forward for improving population estimates, will be discussed in a further paper next year.

Results

Sex ratio patterns in an administrative data source: Patient Register Data (PRD)

There are few administrative data sources that cover the whole population. One data source that has high coverage is the PRD. There are known data quality issues with this data set for purpose of producing population counts; for example, list variation caused at a national level by people moving abroad and not de-registering, and immigrants not registering or registering late.¹⁰ List variation may also be caused by moves within EW, although the unique NHS number given to patients should minimise this risk. These issues may vary by age and sex and this needs to be considered in the following analysis.

Demographic data from patient registers is provided to ONS from two sources. Stock data of all people registered with a GP is provided from National Health Applications and Infrastructure Services (NHAIS). Information on flows is also provided from the Central Health Register



Inquiry System (CHRIS). Both are used in estimating sub-national migration for population statistics purposes. CHRIS data is also provided to the ONS Longitudinal Study.

The following analysis is based on both data sources:

1. the stock of those registered with a GP, provided to ONS for the purpose of calculating sub-national migration (NHAIS).
2. a one per cent sample of flows on to and off of the NHSCR, provided to the ONS Longitudinal Study.

The one per cent flow data is used to examine how the sex ratio pattern in the GP register data is produced. The sex ratio pattern in the GP register is distinct both from a 'natural' population (where only births and deaths determine the pattern) and from the MYEs. Unlike the MYEs, the pattern is not ageing forward over this decade (**Figure 2a**). The sex ratio pattern for the PRD dips around similar young ages to the 2001 MYEs. However, it does not age forward, the dip becomes slightly deeper between 2001 and 2006 (**Figure 2b**) and then recovers to a higher than previous level after age 30.

The sex ratio pattern in the PRD is plotted with the sex ratio pattern in the MYEs for 2001 (**Figure 3a**) and 2007 (**Figure 3b**). Figures 3a and b

demonstrate the different patterns of ageing forward in both datasets. Between 2001 and 2007, from age 16 onwards, the pattern in the MYE moves away from that of the PRD. In 2007 it is not until age 28 that the lines cross and the sex ratio in the PRD exceeds that of the MYE, with a bulging pattern between the ages 35 and 75, until the lines meet again.

Using ONS LS data to understand the sex ratio pattern in the PRD

GP register flows provided for the ONS LS were used to determine why the sex ratio in the PRD data dips and recovers without ageing forward. Preliminary analysis showed that the sample flow data, despite being only just over a one per cent sample, reproduced the distinctive trend in the sex ratio of the national PRD stock data.

Two subsequent analyses were undertaken. Firstly, flows to and from CHRIS were examined. **Figures 4a** and **4b** represent the cumulated flows of entries (re-entries, immigrations) and exits (embarkations, cancellations, deaths) between 2001 and 2005 to and from CHRIS. Cancellations are those people who have failed to respond to letters or

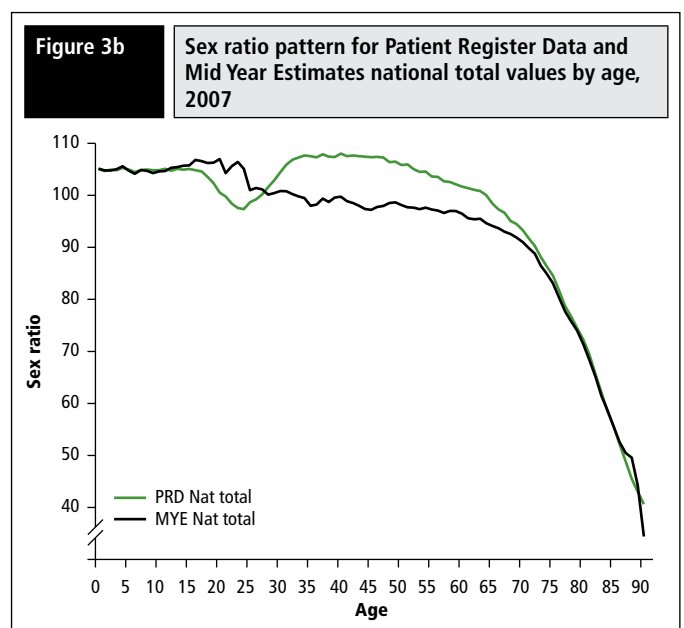
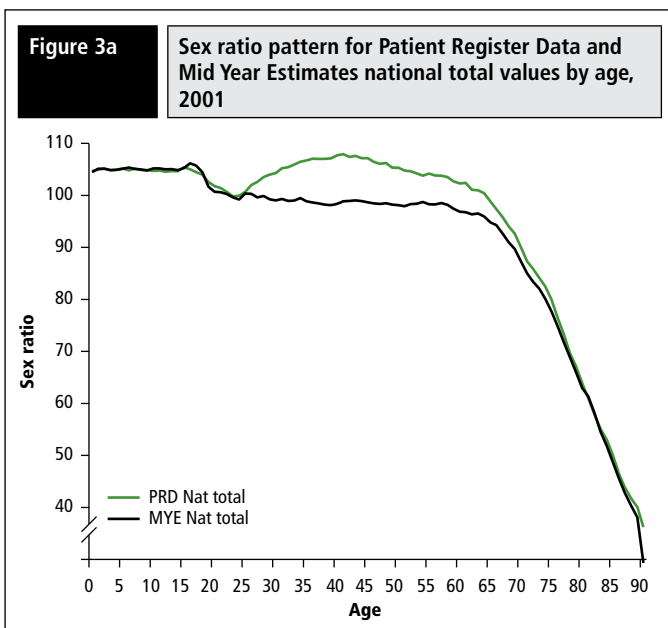
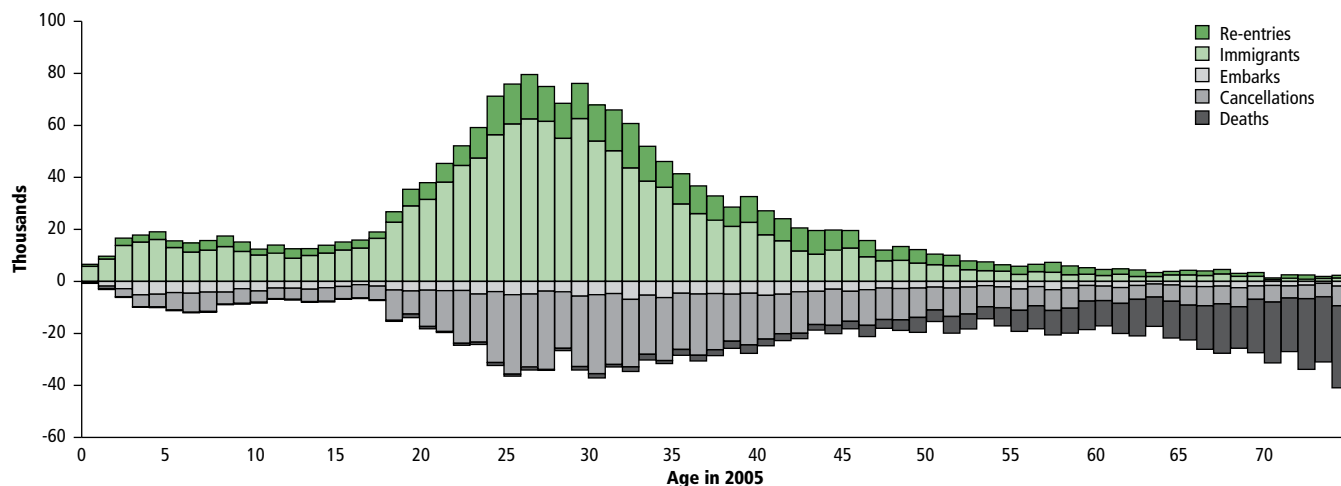


Figure 4a

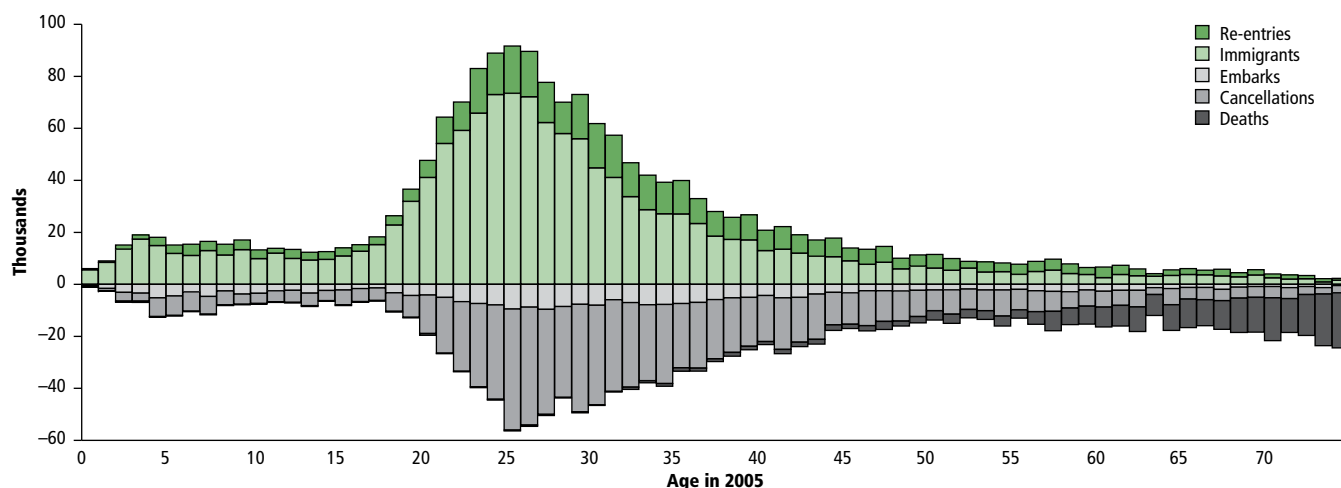
LS sample of flows to and from GP register mid-2001 to mid-2005 rated up to be representative of national flows, Males



Source: ONS Longitudinal Study, sample numbers less than three have been disclosure controlled

Figure 4b

LS sample of flows to and from GP register mid-2001 to mid-2005 rated up to be representative of national flows, Females



Source: ONS Longitudinal Study, sample numbers less than three have been disclosure controlled

have failed to show up for appointments and are cancelled administratively. Embarkations are those who report leaving the country. Cancellations represent a much greater number of the outflows compared to embarkations. Three attributes of the sex imbalances are noteworthy in Figures 4a and b:

- a larger flow of female immigrants (a cumulative maximum of 73 thousand) at an earlier age (24 in 2005) compared to the male population (a cumulative maximum of 62 thousand at age 26 in 2005)
- a larger flow of male immigrants at a later working age compared to the female population (for example at age 30 in 2005 a cumulative figure of 54 thousand for men and 45 thousand for women)
- finally, a much larger outflow of women at ages 30 and over in 2005 compared with men (for example at age 30 the figures are 38 thousand compared with 30 thousand)

Trends at older ages reflect mortality differences by age between men and women.

Further investigations then assessed how many of the cancellations and embarkations among females and males were originally

immigrants. The analysis confirmed that to a large extent the cancellations were of previous immigrants. Both the cumulated flows of the cancellations and embarkations show variability, with both flows increasing at peak migration ages (18 to 40 years). In particular, those who were recorded as being immigrants are responsible for the increase in outflows at peak-migration ages. This observation is more pronounced for the females at peak migration ages (40 thousand at age 26 in 2005 for embarks and cancellations combined as compared to 20 thousand for the males) (Figures 5a and b). This raises the question of whether the sex difference in outflows corresponds to actual migration or whether there is an inherent gender bias in this data source in the tracking of people’s movements. It may also provide evidence that the patient register may not capture migrant outflows of the EW born effectively.

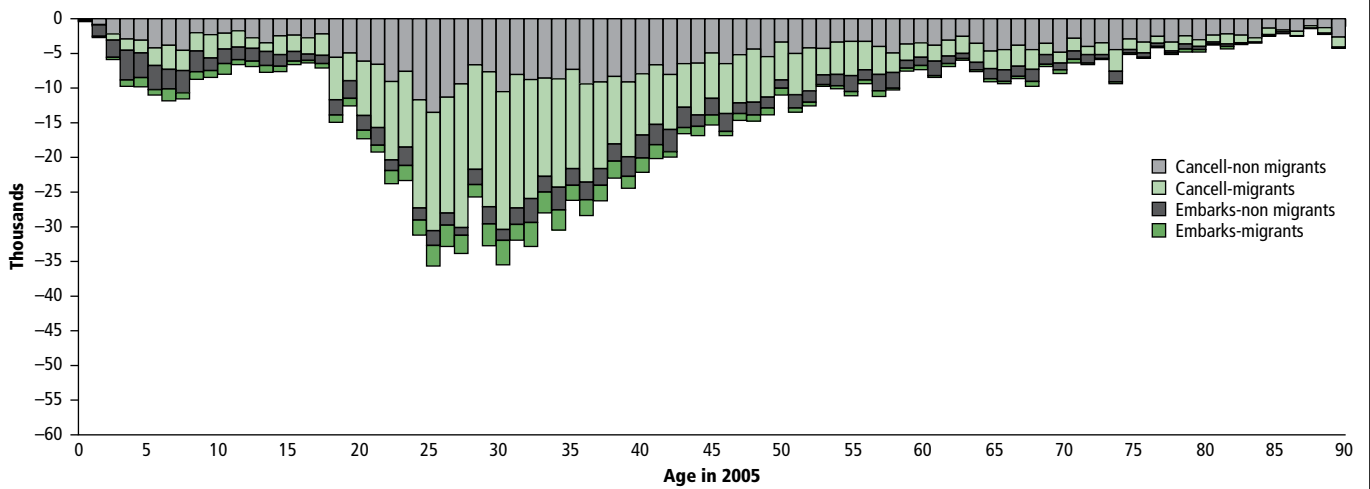
International and UK Countries comparison

International Comparisons

Sex ratios patterns in other countries were examined. In most countries statistical collection of migration data is a by-product of national

Figure 5a

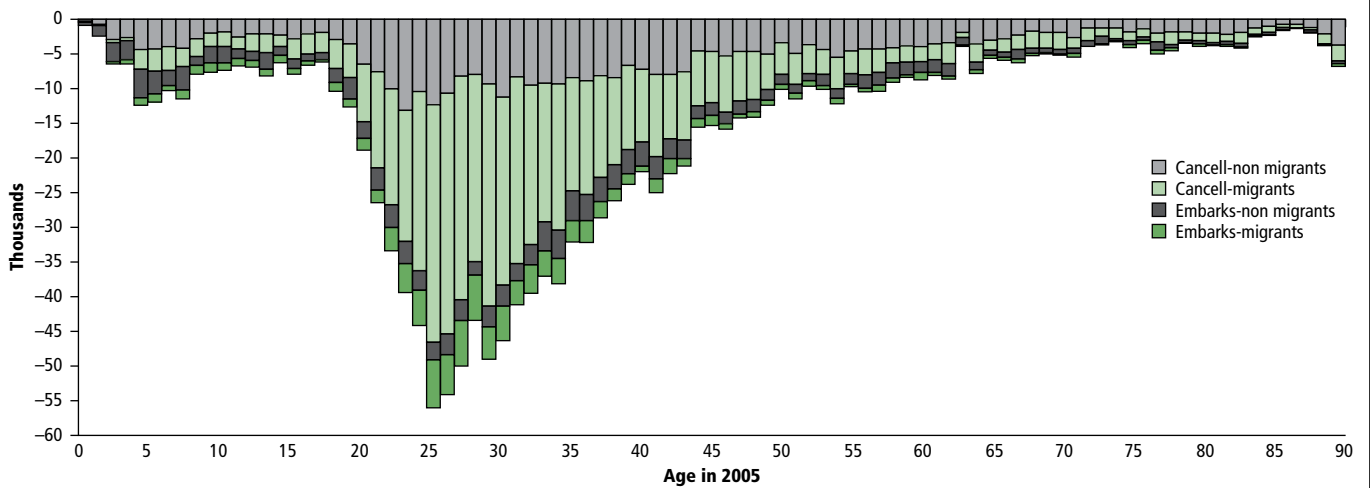
LS data on embarkations and cancellations from GP registers by previous migrant status rated up to be representative of national flows, Males



Source: ONS Longitudinal Study, sample numbers less than three have been disclosure controlled

Figure 5b

LS data on embarkations and cancellations from GP registers by previous migrant status rated up to be representative of national flows, Females



Source: ONS Longitudinal Study, sample numbers less than three have been disclosure controlled

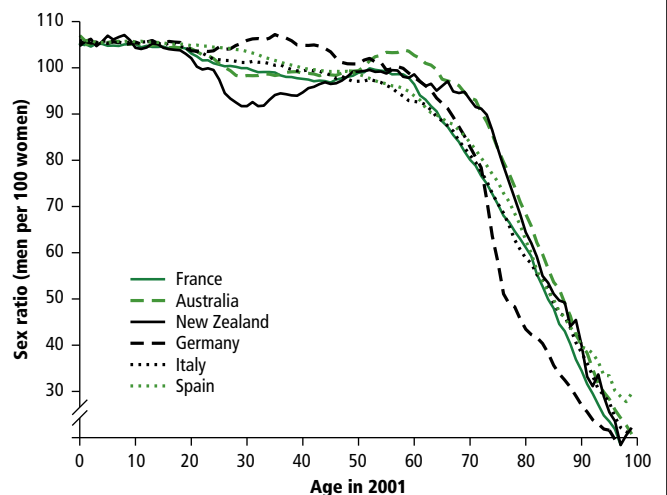
administrative data collection systems. Additionally, countries differ in their definition of a ‘migrant’ or ‘migration’, including deviating from the UN-recommended international migration definitions. Many EU countries have some form of population register. Those that are most reliable fully integrate their register with national administrative data. The Nordic countries have made efforts to improve international comparability through bilateral cooperation and adopting the UN criteria.¹¹ As such, the data for the Nordic countries have the most chance of providing reliable population data and therefore accurate sex ratio patterns.

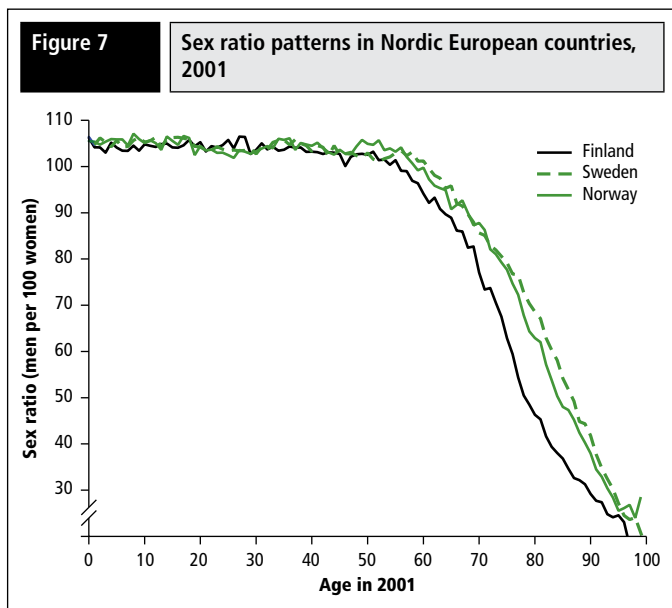
In **Figure 6** and **Figure 7** international data are presented for specific countries, divided into selected OECD countries (Figure 6) and Nordic Countries (Figure 7). The key observations are:

- All OECD countries, except Spain, show some drop in the sex ratio at ages in the 20s and 30s.
- No country has such an abrupt drop in the sex ratio as EW had at 18–19 in 2001
- New Zealand shows the sharpest and deepest dip in sex ratios between ages 20 and 45

Figure 6

Sex ratio patterns in a selection of OECD countries, 2001

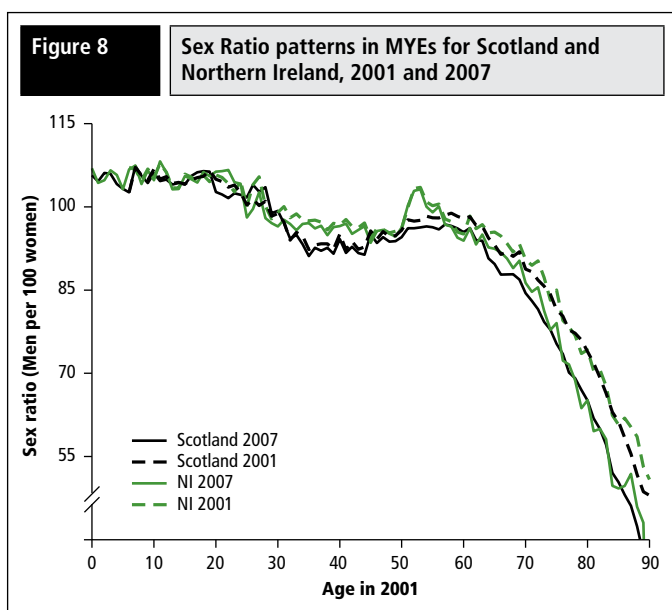




- In general the sex ratio patterns in the Nordic Countries Finland, Sweden and Norway are as would be expected in the absence of any strong differences in migration by sex, with sex ratios not dipping below parity until beyond around age 50 (although Norway shows a slight dip in the sex ratio at ages 25–30).

Sex ratios in Scotland and Northern Ireland (NI)

Although this work is primarily concerned with EW, the sex ratio pattern in Scotland and NI could provide a partial explanation for the sex ratio patterns in EW if the ‘missing’ young EW men were found in those countries. However, drops in the sex ratio were also observed in Scotland and NI (Figure 8). All lines are plotted based upon actual ages in 2001 to show the sex ratio in each of the years effectively preserves that found in 2001. Sex ratios for these countries are more ‘ragged’ because of the smaller population size. Within UK cross border migration also plays a larger role in both countries as compared to EW. The MYEs between 2001 and 2007 simply age forward the sex ratio found in the 2001 Census, except for the peak migration ages where migration has a small effect on the sex ratio, and at older ages where mortality affects the sex ratio.



Does the remaining unexplained difference in 2001 shed light on sex ratios?

Although much of the difference in overall numbers between the 2001 Census and the rolled forward estimates has been explained, the detailed age-sex division of population produces a national sex ratio pattern that has not yet been fully explained. Appendix A (Figure A3) shows the differences remaining by age and sex.

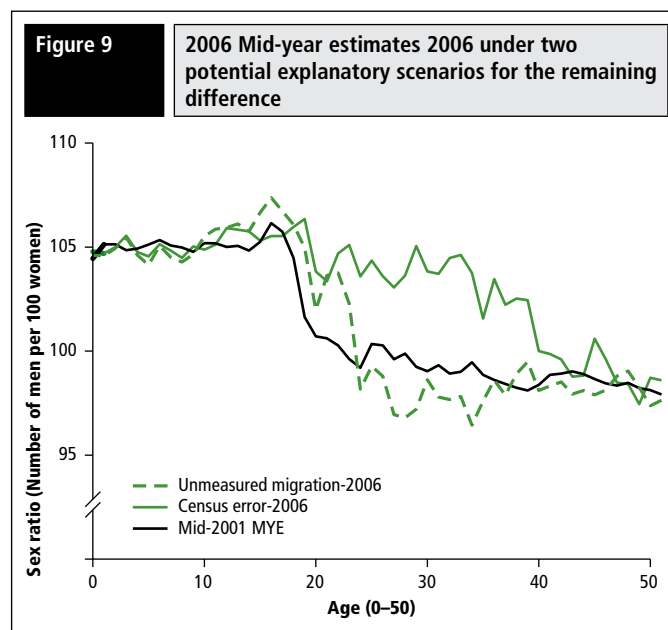
In the following exercise two potential scenarios for attributing the unexplained difference of 209,000 are presented. Both of the scenarios are extremes. The first scenario assumes that the remaining differences between the 2001 Census and the MYEs by age and sex are because the Census was not sufficiently precise in measuring population by age and sex, and that the revisions to the rolled forward estimates created the ‘true’ population. The scenario therefore adds back the net 209,000 population difference (in practice this would involve subtracting from some of the age groups, mainly young females). Under the second scenario, the assumption is the remaining difference at the 2001 Census is unmeasured migration over the 1990s and that similar rates of unexplained migration are occurring evenly each year in the current decade.

Figure 9 shows the sex ratio pattern for both scenarios. Both of the scenarios are likely to be extremes and while neither produces completely acceptable sex ratios they do appear to dampen the unusual effect currently seen in the MYEs. The distinctiveness of the two different scenarios provides further motivation to examine other data sources which could potentially provide a reliable indication of the national sex ratio in EW.

Sex (im)balances in migrant stocks

The final piece of work carried out attempts to account for the type of migrant that is affecting the sex balance in the population. Is it a result of a sex imbalance in overseas born migrants in EW, or in EW born population being abroad? In order to understand sex ratios of migrants and their impact on population estimates, the following analyses were carried out.

A ‘theoretical’ population of those born in EW estimated to be alive today² was derived using birth data and mortality rates only. So no account is taken of migration. The method of calculation is summarised in Box One.



Box one

Calculation of England and Wales born population from births and mortality rates

Since 1837 there has been a legal requirement that every birth that occurred in EW should be registered and therefore a complete series of data is available¹³ relating to the EW born population alive today. Similarly, all deaths (both EW and non-EW born) that occur in EW must be registered. By combining the deaths with information on population data from censuses and, since the 1960s, population estimates, the death figures can be used to estimate a historical series of age and sex specific mortality rates for EW.¹⁴ These can be used with the births data to estimate the number of EW born people alive in 2001 as described below.

$l_{x,y}$ = Numbers alive at age x in year y out of original number of births

$q_{x,y}$ = probability of dying between exact age x and exact age x + 1 in year y

An estimated number of people alive at exact age 1 is produced by first taking births to produce a population age 1 one year on.

$$l_{1,y+1} = l_{0,y} - l_{0,y} \times q_{0,y}$$

Subsequent populations in the cohort are then produced such that

$$l_{x,y+x} = l_{x-1,y+x-1} - l_{x-1,y+x-1} \times q_{x-1,y+x-1}$$

In order to produce the average number alive in the interval between two exact ages x and x + n (${}_nL_x$) four adjacent populations are averaged. For example,

$$L_{20,2001} = (l_{20,2000} + l_{20,2001} + l_{21,2001} + l_{21,2002})/4$$

Effectively the average of the 1980 cohort surviving at exact ages 20 and 21 and the 1981 cohort at exact ages 20 and 21.

The population of EW estimated to be alive was then compared with both the EW born populations in the 2001, 1991 and 1981 Censuses, and, the EW born in the population estimates for the same years (derived using the census proportions). The difference between this ‘theoretical’ population and the Census (or estimated MYE) stocks of EW born residents in EW gives an estimate of the stock of EW born emigrants.

Conversely the stock of foreign born immigrants can be obtained directly from census counts (which can also be used to estimate the stock of foreign born in the mid-year estimates). The patterns in sex ratios among the derived stocks of EW and non-EW born migrants are shown in **Table 1**.

Differences between the ‘theoretical’ population derived from births and mortality rates and other population estimates (census or MYEs) could be attributed to a number of factors other than migration: mortality calculations; errors in the census; and, for estimates, assumptions made about country of birth, and past errors in adjusting population estimates after previous censuses. Analyses of sensitivity using different assumptions showed that only a small proportion of the difference could result from mortality calculations, at least at younger ages, and that only extreme assumptions for country of birth would affect the conclusions.¹² It is therefore assumed that migration does account for the difference.

Table 1

Sex ratios for the Census and mid-year estimates 1981, 1991 and 2001 for the stock of non-England and Wales immigrants and for England and Wales born emigrants

Age	Census			Latest mid year estimates		
	1981	1991	2001	1981	1991	2001
	Non-EW born resident in EW (based on counts in 1981 and 1991, 2001 adj. for under enumeration)			Non-EW born resident in EW (Estimated using census data)		
15–19*	106.6	102.2	106.7	107.1	103.0	104.2
20–24	98.0	87.4	87.5	98.0	90.8	86.0
25–29	98.0	93.4	87.4	98.5	95.5	91.1
30–34	94.0	96.5	93.4	94.5	97.1	96.9
35–39	97.4	95.3	95.5	98.0	95.8	96.1
40–44	106.2	90.5	95.2	106.7	90.9	95.5
45–49		94.1	91.7		94.3	91.9
50–54		103.6	86.4		103.5	87.3
55–59			91.4			90.3
60–64			96.6			96.9
	Estimated EW born emigrants			Estimated EW born emigrants		
15–19*	113.2	121.4	112.4	110.2	110.8	108.9
20–24	117.8	153.2	149.5	125.8	128.1	149.1
25–29	123.3	156.1	192.6	124.5	154.1	131.4
30–34	126.2	148.5	180.5	123.7	152.5	140.4
35–39	116.0	143.2	156.5	114.1	140.1	154.3
40–44	115.1	138.1	143.8	110.7	134.0	142.5
45–49		116.1	137.9		116.6	134.1
50–54		109.8	130.0		110.2	127.8
55–59			112.4			113.8
60–64			103.1			103.6

* Figure for 16–19 shown for 1981

The analysis in Table 1 brings together the sex ratios among migrant stocks (non-EW born residents in EW, and EW born residents who are abroad) derived from the 2001 Census, MYEs and the ‘theoretical’ population. The table is divided in two sections with the first (left) section showing the sex ratios based on the three censuses. Section two (right) is based on the latest MYEs, including the adjustments made after the 2001 Census. The table concentrates on the peak migration ages, although data beyond the peak migration ages are provided for 1991 and 2001 so that the cohort patterns can be followed.

The rows show age at the time of the census/MYEs. Cohorts can be followed in the shaded diagonal in the table; those aged 15–19 in 1981 are 10 years older by 1991 and 20 years older by 2001. The top part of the table shows sex ratios among the stock of non-EW born immigrants, showing slightly more women than men (sex ratio under 100), whereas the bottom part of the table shows the sex ratios among the EW born emigrant stock are male dominated. The fact that this trend can be observed in all three census years from ages above 19 warrants further investigation.

The estimated sex ratios for EW-born emigrant stock aged 25–39 in 2001 for the 2001 Census results are greater than 150, and approach 200 for the 25–29 group (left three columns of Table 1). The revisions and addition of men through the ONS Longitudinal Study (ONS LS)

adjustment considerably tempered these extreme sex ratios among emigrant-stock derived from the MYEs (right three columns in Table 1), with only the sex ratio for 35–39 years olds remaining over 150. This analysis can be seen as providing additional support for adjustments made.

The patterns here suggest that the fall in sex ratios at adult ages is driven mainly by much greater net outflow of EW born men than women although there are also slightly more females than males from abroad in the population. Thus, when examining migration flows particular attention needs to be drawn to the age and sex flows of the EW born as well as the flows of the non-EW born.

The numbers behind the information in **Table 1** can be found in an earlier paper.¹⁵

Discussion

How is the sex ratio pattern in the MYEs to be explained?

The sex ratio in the mid-2001 population estimates predominantly reflects the 2001 Census results. In many countries, low sex ratios at working and migratory ages have been observed in census counts. Evidence has shown¹⁶ that virtually all censuses struggle with the problem of undercounting of young men; it is assumed that more mobile young males particularly in the age range 20 to 34 are likely to be undercounted compared to their female counterparts.¹⁷ The increased proportion of women in the immigrant stock is likely to reflect a slight sex imbalance in the non-EW born migrant flows.

The international analysis confirms this by showing a dearth of young men in many developed countries' censuses and population registers. To establish whether under-count is universal, or if young men were simply moving between countries with consequently some countries showing raised sex ratios, extensive and careful analysis of both males and females by country of birth across all countries of the world would be required. For many countries such data is simply not available. Attempts have been made to estimate the stock of the EW born population abroad by Redfern¹⁸ using the sparse data that is available, although others have challenged his model.¹⁹

Findings regarding data quality in the Census also raise the question of how censuses define residence, an area already identified as of great importance for the 2011 Census.²⁰ One of the great difficulties in any census is measuring the resident population for whom the concept of usual residence is ambiguous. The 2001 Census definition of 'usual resident' was '...someone who spends the majority of their time residing at that address'.²⁰ Applying this definition may have been problematic for certain sub-groups, such as those who have no usual residence anywhere in the UK, or those who are resident in the UK but do not have a usual residence in any one place here. These were more likely to be people who were single or not part of a traditional household and who consequently may not have been 'captured' by the census, including in the Census Coverage Survey. An additional definitional issue is that a long-term migrant in the 2001 Census is someone who has left their 'usual residence' for longer than six months, whereas the cut-off point for the MYEs is 12 months. If individuals under both definitions were disproportionately male then this would have impacted on the sex ratio.

Trends in the PRD sex ratio have been shown here to be driven primarily by the registration and cancellation of female migrants. Females are reported to be more compliant with administrative requirements and to register with a GP on arrival at the new location. Other evidence from literature suggested that the dip in the sex ratio around the younger ages 18 to 24 was explained by men making use of the doctor infrequently and therefore being less likely to be on the PRD.²¹ However, for this to be

true, young EW-born men would have to be de-registered (cancelled) first from the PRD. Evidence does not seem to support that there were many more cancellations among males compared to females. The 'bulging' profile in the PRD (Figure 2a and b) above the age of 30 appears to be the result of both higher numbers of males (migrants) registering at a GP practice and an excess outflow (cancellations and embarkations) of females compared to males, observed in the ONS LS data.

Although the accounting exercise analysing the effect of the 209,000 unexplained differences between rolled forward MYEs and the Census on sex ratios provided more plausible patterns under both scenarios, both are likely to be extremes and offer different sex ratio patterns. The first scenario, where the revised rolled forward estimates are accepted, appears to provide the more plausible sex ratio in that it is more consistent with the sex ratio from register based countries and does not show a sharp change at a particular age.

Issues addressed in order to prepare for the 2011 Census and beyond

Improved management of the 2011 Census including capturing through the Census Coverage Survey, and experience drawn from previous censuses, will contribute to minimising of any future under-estimation of specific age and sex groups. There have also been improvements to the information provided to respondents on the definition of those who should complete the census to ensure more complete coverage. Additionally, new questions will be asked on intended length of stay and second residence to allow a more flexible series of outputs using different residence definitions, including a greater compatibility with the MYE definition.

Work is still in progress to understand why the sex ratio pattern ages forward over the intercensal period. If the dip in the sex ratio has been caused by the under-enumeration of young males in 2001 then improved enumeration may give a different pattern in 2011. However, questions remain whether sufficient measures are in place to capture returning short term migrants abroad at the time of the Census. Extensive work on migration statistics after the 2001 Census has focussed on the geographical distribution of the population, with improvements to the measurement of the distribution of international immigrants and emigrants, with further improvements proposed including those of student moves subnationally. Improvements in the International Passenger Survey (IPS) will also see larger sample sizes and therefore greater precision, although they will not necessarily change the overall totals and age sex patterns. Further, information about the accuracy of sex ratios in net migration data by detailed age group is still limited and provides an important area for future work.

One contingency for the coverage estimation for the 2011 Census will be the use of target sex ratios. After the 2011 Census, target sex ratio could be used in Census coverage adjustments as a target control. They could be used, if evidence suggested that the Census was incorrect, to rebase from Census to estimates to minimise the gap between the rolled forward estimates from 2001 and the 2011 Census. They could also potentially be used to adjust future mid-year estimates.

Future work

At a national level there appears to be evidence from the PRD and from the remaining unexplained difference between the 2001 Census and rolled forward estimates, warranting examination of the age sex patterns of migrants. Improvement to sex ratios in population statistics would enhance confidence in those products. The methodology for putting together the various sources used for the migration element of the national population estimates by age and sex contains a number of assumptions that will be explored further.

At the moment, it is not possible to draw any strong conclusions from the analysis provided here of what the 'true' sex ratio pattern in EW should be given the two sources examined (MYEs – which is predominantly census driven – and PRD). In order to make this judgement, further work should be undertaken and a third administrative data source would assist. The Department for Work and Pension Longitudinal Study (WPLS) may prove a useful source to be examined to shed light on the sex ratio patterns of EW. The aim is to develop target sex ratios based on the triangulation of data sources (MYEs, PRD and WPLS data) at the national level. Further international analyses using data from other countries may provide some explanation for the sex ratios at the national level in EW, or at least help in providing an indication of plausible patterns.

This article has focused on the national sex ratio pattern; however there is also interest in sub-national sex ratios. At the sub-national level errors in sex ratios are as much, if not more, likely to result from issues with the computation of sub-national migration from GP register data, as well as the distribution of international immigration and emigration. If men are generally worse at registering with a doctor then this may produce a sex biased set of sub-national migration figures. At the subnational level, target sex ratios could be developed based on typologies of LAs or clustering of areas with similar sex ratio patterns.

Conclusion

This article has summarised work contributing to an explanation of the sex ratio pattern seen in population estimates since 2001. Questions remain as to the source of any error in the observed sex ratio pattern. For the 2011 Census there will be even better management and capturing in the CCS and lessons drawn from previous censuses will minimise both under- and over counts in the 2011 Census. Also improvements have been made to definitions of who should complete the survey in order to improve coverage, and extensive work has been undertaken to improve migration figures at national and subnational level.

Further research will aim to increase understanding of the relative importance of definitional issues and population numbers of males and females, and/or whether assumptions used in age distributions of migrants are pertinent. At this stage, there is no compelling evidence to adjust the population at the national level, but a number of further avenues such as the development of target sex ratios, have been identified that would prepare for the 2011 Census and beyond.

At the subnational level other improvements to international and internal migrants could potentially have an effect on the sex ratios of some local authorities. Any examination of subnational sex ratios needs to wait until after these are implemented.

Key findings

- Sex ratio patterns in mid-year estimates in this intercensal decade produce a pattern that requires explanation
- Sex ratio patterns in patient register data differ from those in mid-year estimates
- There is evidence that the national sex ratio pattern for patient registers is driven by greater flows of women on to and off of the patient registers
- Part of the explanation for the mid-year estimates' sex ratio pattern may lay in the age sex pattern of the remaining 209,000 difference between rolled forward estimates and the 2001 Census

Acknowledgements

The authors wish to acknowledge the careful and patient help of the ONS Longitudinal Studies team, in particular Kevin Lynch, and also a discussion with Prof Mike Murphy on international comparisons. The authors also acknowledge the comments of colleagues in ONSCD and of reviewers of an earlier version of this article.

References

- 1 In less developed countries sex ratio may drop at young ages as infant mortality is higher among boys and in some countries, sex ratios at birth can be biased (normally towards male births) when a strong sex preference exists and selective abortion takes place.
- 2 Shryock H, Siegel J, *et al* (1976) *The Methods and Materials of Demography*. Condensed edition by Edward G Stockwell, Academic Press.
- 3 Pereira R (2002) 'The Census Coverage Survey – The Key Element of a One Number Census'. *Population Trends* 108 pp. 16–23. Available at: www.statistics.gov.uk/downloads/theme_population/PT108.pdf
- 4 Available at: www.statistics.gov.uk/census2001/onc_avec_eval.asp#assessment
- 5 Comparison of mid-2001 population estimates – Implications of 2001 Census for mid-year population estimates. Available at: www.statistics.gov.uk/populationestimates/01/commentary.asp
- 6 Office for National Statistics (2004) 2001 Census: Manchester and Westminster Matching Studies, Summary Report. Available at: www.statistics.gov.uk/downloads/theme_population/2001CENSUSM&WMATCHINGSUMMARY.pdf
- 7 Office for National Statistics (2004) 2001 Census Local Authority Population Studies: Full report. Available at: www.statistics.gov.uk/downloads/theme_population/LAStudy_FullReport.pdf
- 8 Office for National Statistics (2003) Revisions to Population estimates-Methodology. Available at: www.statistics.gov.uk/about/methodology_by_theme/Revisions_to_Population_Estimates/downloads/Methodology_for_revision_to_mid-2001.pdf
- 9 Office for National Statistics (2004) Local Authority studies – The Longitudinal Based Study – Consequential Adjustment. Available at: www.statistics.gov.uk/downloads/theme_population/LAStudy_LS_ConsequentialAdjustment.pdf
- 10 Diamond I, Abbott O, Jackson N (2003) 'Key issues in the quality assurance of the One Number Census'. *Population Trends* 113 pp. 11–19.
- 11 Poulin M and Perrin N (2003) Can UN migration recommendations be met in Europe? Available at: www.migrationinformation.com/Feature/print.cfm?ID=139
- 12 To move from Census to the MYE an assumption has to be made about the division by country of birth. In this analysis it is assumed that the population difference (by age group and sex) between Census and MYEs is divided proportionately by the proportions found in the Census. Only when the division was assumed to be very substantially different from the Census and the division was assumed to vary substantially by sex were the results of the analysis materially affected.
- 13 Office for National Statistics (2007) Birth Statistics. Series FM1 no. 35, ISSN 0140-2587 pp. xiii. Available at: www.statistics.gov.uk/statbase/Product.asp?vlnk=5768
- 14 These rates were produced by the Government Actuary's Department, and are now available for research purposes, on request, from the Mortality team in the ONS Centre for Demography.
- 15 Smallwood S and De Broe S (2008) Unravelling sex ratios, what can they tell us about the Census, population estimates and recent trends in migration? Paper presented at the EAPS conference in Barcelona 10–12 July 2008. Available at: <http://epc2008.princeton.edu/download.aspx?submissionId=80292>
- 16 Phillips H, Anderson B, Tsebe N (2003) 'Sex Ratios in South African Census Data, 1970–96'. *Development Southern Africa* 20(3) pp. 387–404

and: Bycroft, C (2007) 'Challenges in Estimating Populations'. *New Zealand Population Review* 32 (2):21-47.

- 17 Research at the Office for National Statistics (2003) on migration data found some evidence that emigration had been underestimated in the 1990s and that the majority of the underestimation is likely to be of males. However, the underestimation did not account for all of the observed fall in the sex ratio, in particular for the younger age group 25-29.
- 18 Redfern P (2001) 'A Bayesian Model for Estimating Census Undercount, Taking Emigration Data from Foreign Censuses'. *International Statistical Review*, 69, 277-301 and: Redfern, P (2003) 'Estimating Census Undercount by Demographic Analysis: New Approaches to the Emigrant Component'. *Journal of Official Statistics*, vol 19, pp. 421-448.
- 19 Discussion on the meeting 'The 2001 census and beyond' *Journal of Royal Statistical Society Series A* (2004) vol 167, Part 2, pp. 229-248. See comments by Len Cook (Office for National Statistics, London) and James Brown (University of Southampton).
- 20 Smith C, Jefferies J (2006) Population bases and statistical provision: towards a more flexible future? *Population Trends* 124 pp. 18-25.
- 21 Scott A and Kilbey T (1999) 'Can Patient Registers Give an Improved Measure of Internal Migration in England and Wales?' *Population Trends* 96, pp. 51.

Appendix A

This article discusses the patterns of sex ratios, but the ratios themselves are derived from the numbers of males and females by age and sex. Charts are presented below on the population numbers in the MYEs and the PRD for the selected years 2001 and 2006, from which the sex ratios in this article have been derived (Figures A1 and A2).

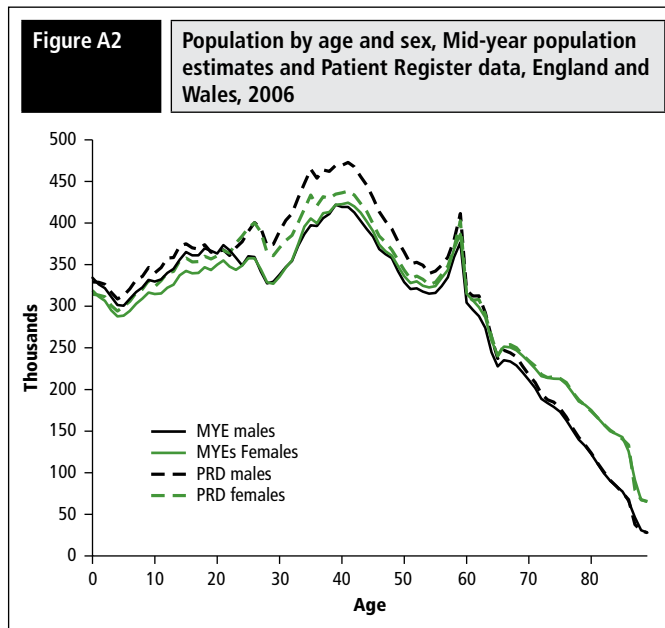
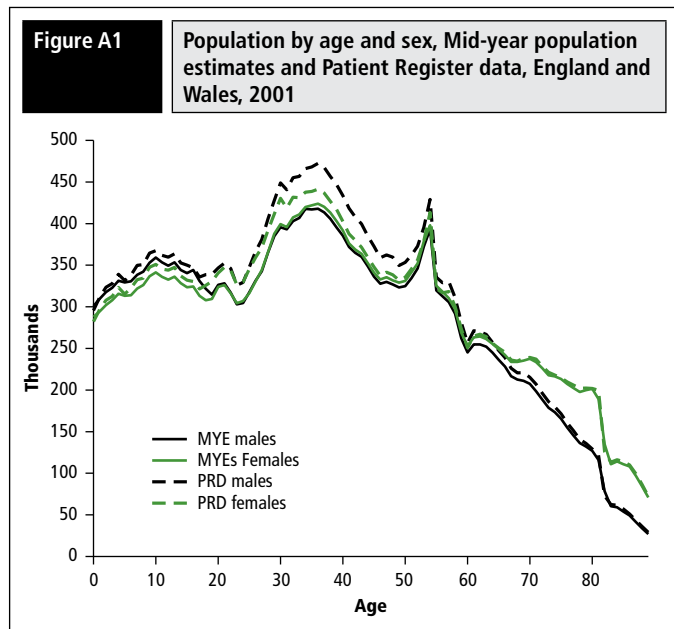


Figure A3 shows the remaining unexplained difference of 209,000 between the rolled forward mid-year estimates and the 2001 Census. This shows that there is still an unexplained difference for men aged in their early 20s and for men in their late 20s through to their 40s. For women aged 20 to 35 there were actually more women found than expected in the census, conversely there were fewer women found aged 35 to 45. The net difference for men is 238,000 and for women -29,000.

