

Identifying inter-censal drift between 1991 and 2007 in population estimates for England and Wales

Sofie De Broe, Nicola Tromans, Steve Smallwood, Julie Jefferies

Note: this paper is work in progress and constitutes a version of the intermediate paper

1. Introduction

One of ONS Centre for Demography's key challenges is to reduce intercensal drift in population estimates. This adds credibility and confidence to the estimates. Inter-censal drift of population measures is the accumulation of error in the estimation process. Over time this affects the size and distribution of the population estimates geographically.

The extent to which population drift has occurred since 2001 will not be known until results from the 2011 Census are available. However, inter-censal drift during the 1990's can be observed when comparing the population estimates based on the 1991 Census with those based on the 2001 Census. This research explores how inter-censal population drift can be observed in demographic measures, such as fertility rates, sex ratios and mortality rates. This is measured by observing the distribution of these measures by geographical area. Evidence so far for the 1990's suggests that the estimation error increases the spread of these measures over time.

This paper will then consider lessons to be drawn both for the quality of inter-censal estimates and for the rebasing to the 2011 Census.

This research has used the following series and inter-censal drift estimates:

- i) 1991-2001 demographic measures using population estimates based on the 1991 Census
- ii) 1991-2007 demographic measures using population estimates based on the 2001 Census
- iii) Inter-censal drift estimates in 2001: difference between the rolled forward (1991 based) population estimates for 2001, and the 2001 population estimates (2001 based).

For the Total Fertility Rate (TFR): inter-censal drift estimated for females aged 15-44

For the sex ratio: population estimates for the 20-29 years olds. National sex ratios for all ages hide substantial variation between age groups. Because a dip in the sex ratio pattern (Figure A in Annex) occurs at the younger age groups, the research has focussed on that age group.

For mortality: life expectancy-this analysis still has to be completed

2. Exploring signs of inter-censal drift: comparing 1991 based and 2001 based series

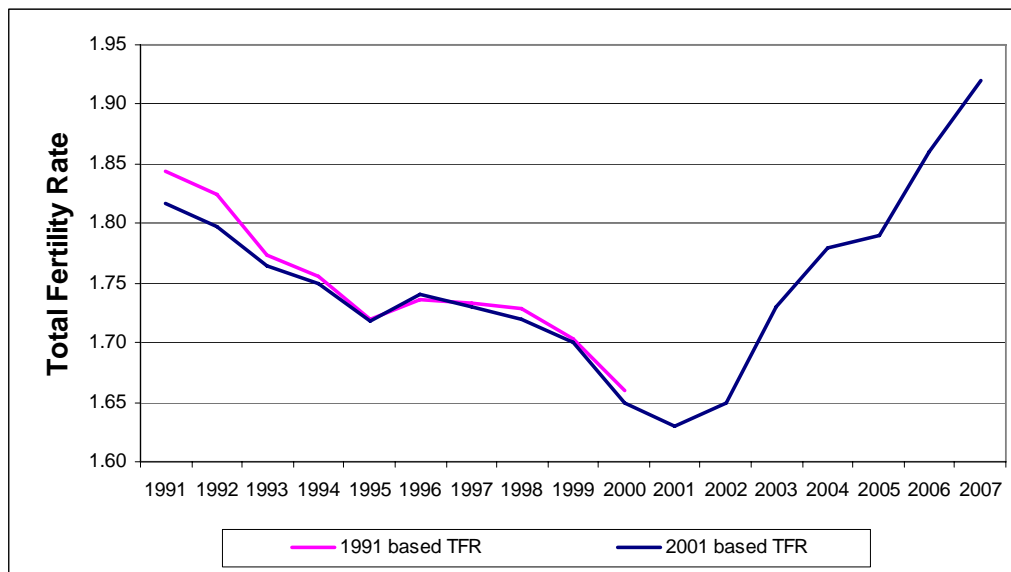
2.1 Differences between the 1991 based and the 2001 based demographic measures

The differences between the 1991 based and the 2001 based demographic measures include both the differences which have been attributed (i.e. over-estimation of the mid-1991 population and international migration mis-estimation) and the intercensal differences which still remain (unattributable population change).

2.1.1 TFR

1991 based population estimates for England and Wales were revised in light of the 2001 Census, which resulted in changes to the TFR. Figure 1 compares the 1991 based TFR series for England and Wales with the 2001 based TFR series. Absolute differences between the two series at the national level were generally between 0.01 and 0 but in 1991 and 1992 the difference was slightly larger: 0.03.

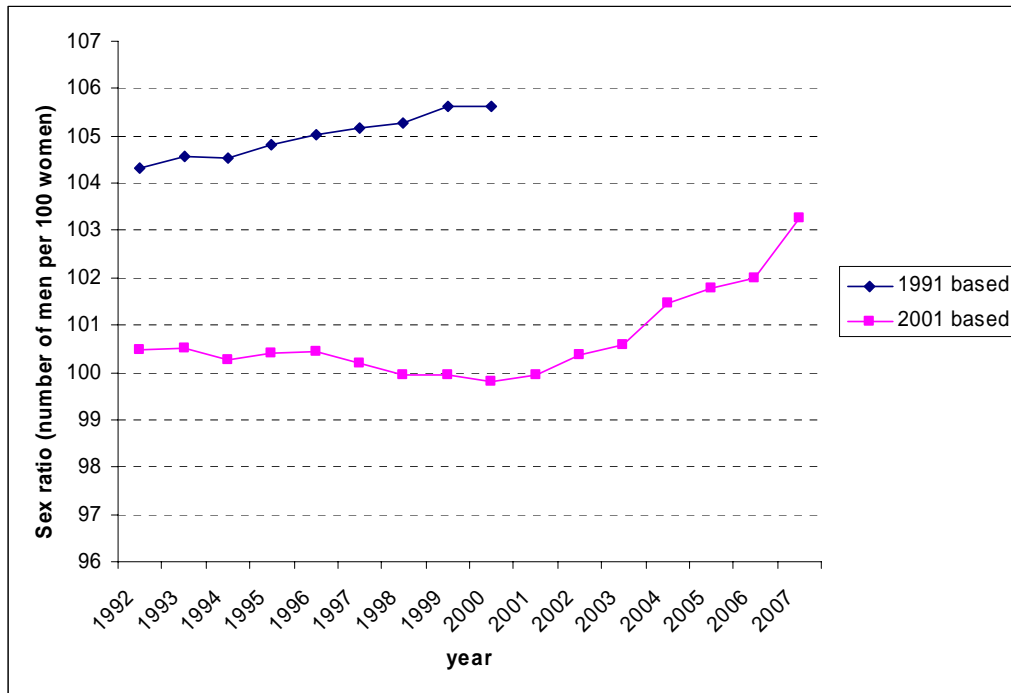
Figure 1: Total Fertility Rate, England and Wales, 1991-2007.



2.1.2 Sex ratio

1991 based population estimates for England and Wales were revised in light of the 2001 Census, which resulted in changes to the sex ratio. Figure 2 compares the 1991 based sex ratio for England and Wales with the 2001 based sex ratio for 20-29 years olds. Absolute differences between the two trends at the national level are generally large because of the revisions that were made post the 2001 Census.

Figure 2: Sex ratio patterns for 20-29 year olds England and Wales, 1991 and 2001 Census based



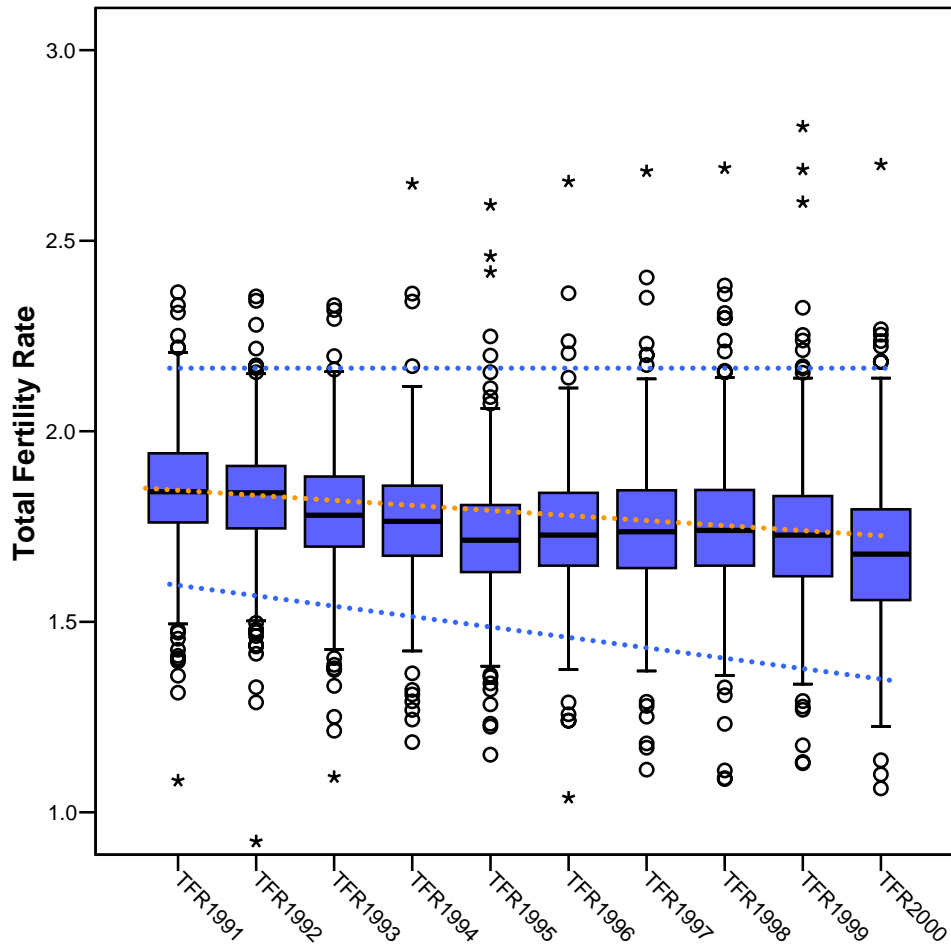
2.2 Distribution of the 1991 based measures

2.2.1 TFR

The distribution of the 1991 based and 2001 based TFRs were compared to illustrate the impact that population drift can have on the TFR. Figure 3 illustrates the distribution of 1991 based local authority TFRs. Areas where the population is very small are clearly susceptible to large changes in TFR between consecutive years e.g. Isles of Scilly and City of London have TFRs which are classed as outliers with a very low TFR and also a very high TFR in different years between 1991 and 2000. In figure 3, the inter-quartile range widens slightly beyond 1995 (boxes increase slightly in length). The local authorities which fall above the upper quartile and below the lower quartile also begin to lie further from the median beyond 1995 (gradual increase in span length of whiskers and the outliers).

Some outliers in figure 3 can be explained: Durham, Oxford and Cambridge, for example, have experienced an increase in the proportion of students and graduates in their populations. Students have lower fertility while studying than women of the same age who are not in higher education and graduates tend to enter motherhood later than women without higher qualifications. This postponement effect for students has had a downwards impact on the TFR in these areas as the size of the student population relative to the non-student population has increased.

Figure 3: Distribution of the 1991 based TFRs, Local Authorities within England and Wales 1991-2000



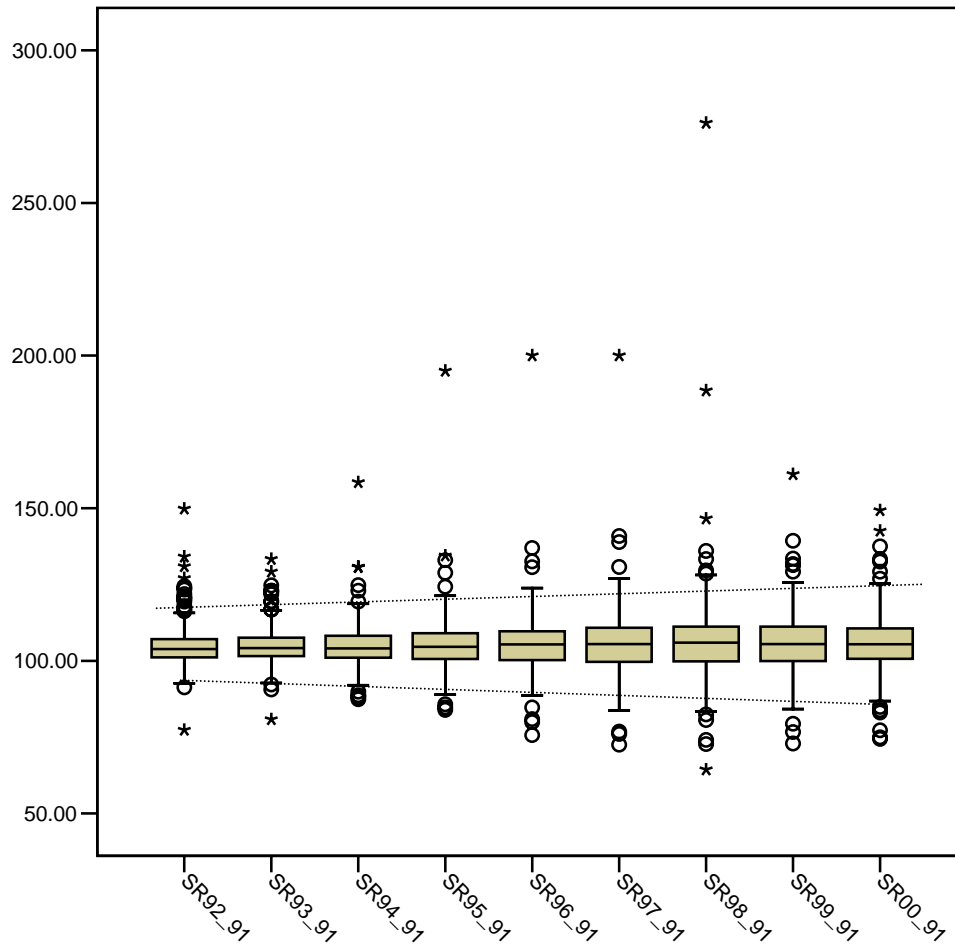
Note: Isles of Scilly is excluded for 1996-1998 and 2000 to enable a sensible scaling to be used.

2.2.2 Sex ratio

The distributions of the 1991 based sex ratios for 20-29 year olds are presented to investigate the potential impact of intercensal drift on the sex ratio at local authority level. The box plots for the 1991 based sex ratio (Figure 4) show:

- local authorities with small populations are likely to be outliers (City of London, Isles of Scilly, Richmondshire)
- local authorities with small populations (City of London) are susceptible to large changes between consecutive years-outliers on either side of the boxplot (minimum and maximum)
- interquartile ranges widen between 1992 and 2000-the main part of the sex ratio distribution- is becoming more widespread
- outliers are lying increasingly further away from the end (whiskers) of the boxplots between 1992 and 2000
- narrowing of the distribution between 1999 and 2000

Figure 4: Distribution of the 1991-based sex ratio for 20-29 year olds, local authorities for England and Wales, 1992-2000



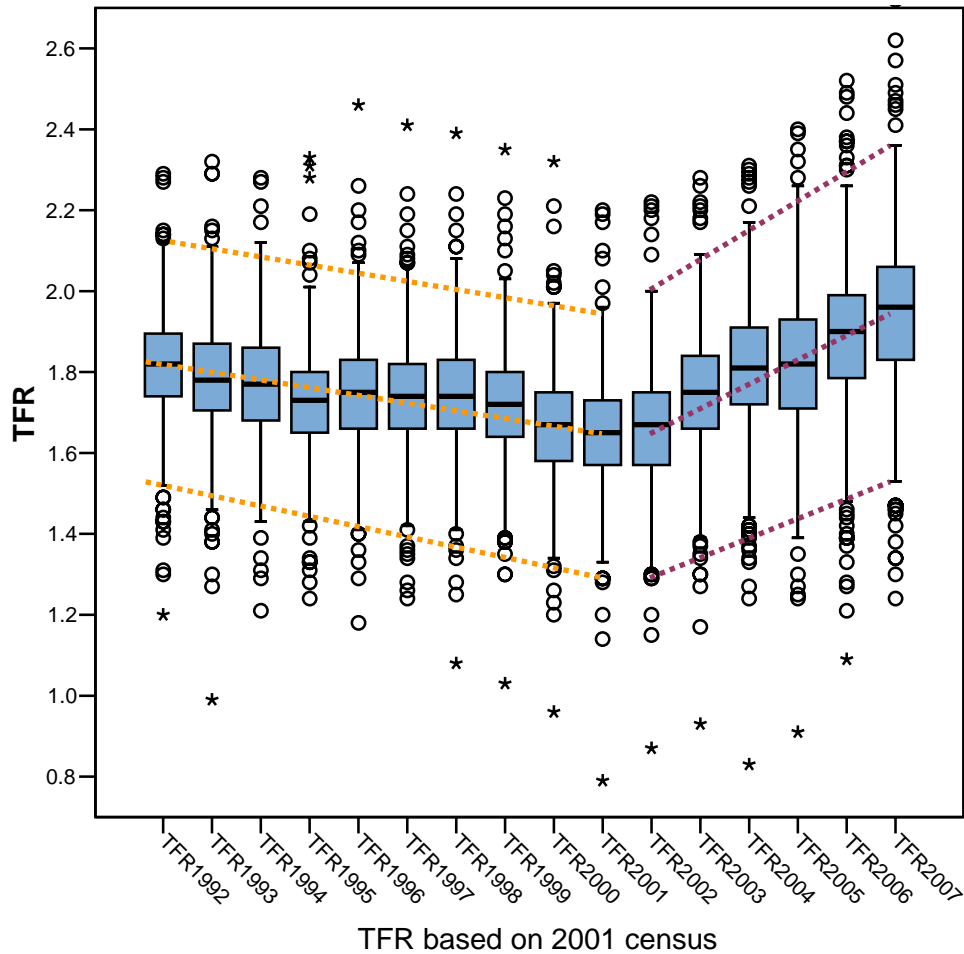
2.3 Distribution of the 2001 based measures

2.3.1 TFR

Figure 5 shows that between 1992 and 2000, the spread of the main part of the 2001 based TFR distribution (box and whiskers) remains roughly constant (orange lines linking median values, upper and lower quartiles stay roughly parallel).

Between 2001 and 2007 however, the spread of the 2001 based TFR distribution increases (purple lines linking the upper whisker points and the lower whisker points diverge from the line linking the median values). The increasing length of the boxes and the whiskers and increasing number of outliers beyond 2002 could be a sign that population drift has occurred in the current inter-censal period since 2001.

Figure 5: Distribution of the 2001 based TFRs, Local Authorities within England and Wales 1992-2007

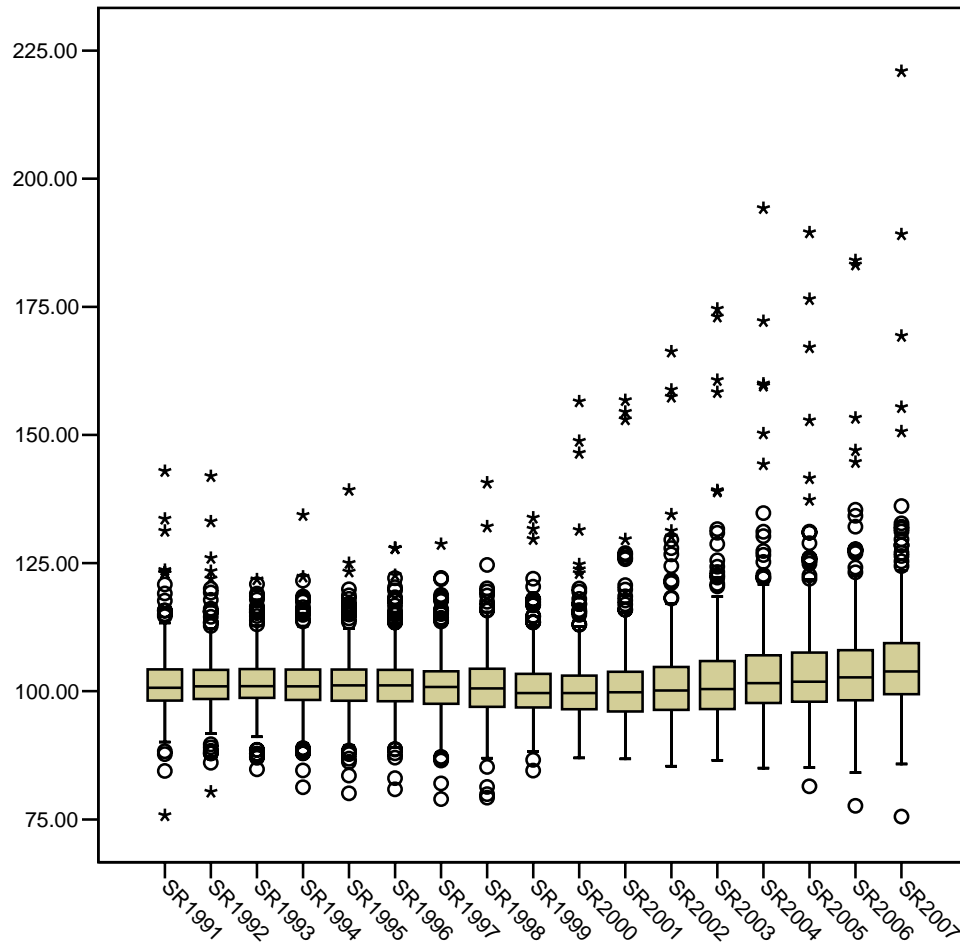


2.3.2 Sex ratio

The box plots for the 2001 based sex ratio (Figure 6) show:

- the pattern of the boxplots reflects the declining sex ratio values until 2000 and increasing sex ratio values until 2007
- between 1991 and 2001 the spread of the main part of the sex ratio is roughly constant possibly because the post-2001 Census revisions addressed intercensal drift between 1991 and 2001
- spread of the sex ratio distribution increases between 2001 and 2007 - the boxplots show a widening of the interquartile range and a lengthening of the whiskers from 2001 onwards
- widening of the interquartile range predominantly happens at the top of the distribution of the sex ratio
- outliers' values becoming increasingly extreme from 2001 onwards for the higher values in the sex ratio

Figure 6: Distribution of the 2001-based sex ratio for 20-29 year olds, local authorities for England and Wales, 1991-2007

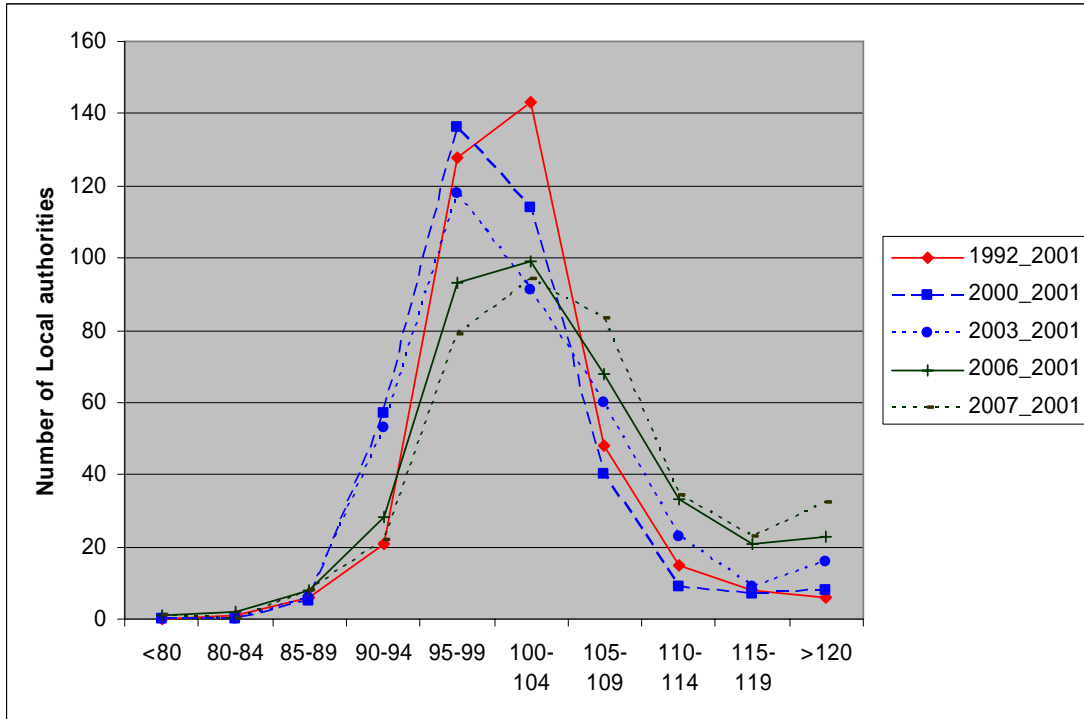


Figures 3 and 6 suggest that inter-censal drift may be occurring over the period 2001-2007. Intercensal drift is reflected in the shape of the distribution of the sex ratio.

Figure 7 shows the frequency distribution of the sex ratio based on the 2001 Census at local authority level. The curves peak at the modal sex ratio range. The modal range in 2000 and 2003 is lower compared to earlier years reflecting the decline in the sex ratio, as shown in Figure 2. The plot also shows that the sex ratio profile becomes more dispersed around the modal sex ratio range as time since the last census increases:

- in 1992, 324 local authorities fell between a 95 and 109 sex ratio
- in 2000, 307 local authorities fell between a 90 and 104 sex ratio
- in 2003, 269 local authorities fell between a 95 and 109 sex ratio
- in 2006, 260 local authorities fell between a 95 and 109 sex ratio
- in 2007, 256 local authorities fell between a 95 and 109 sex ratio

Figure 7: Sex ratio profile for 20-29 years olds, local authorities within England and Wales 1992, 2000, 2003, 2006 and 2007 based on the 2001 Census



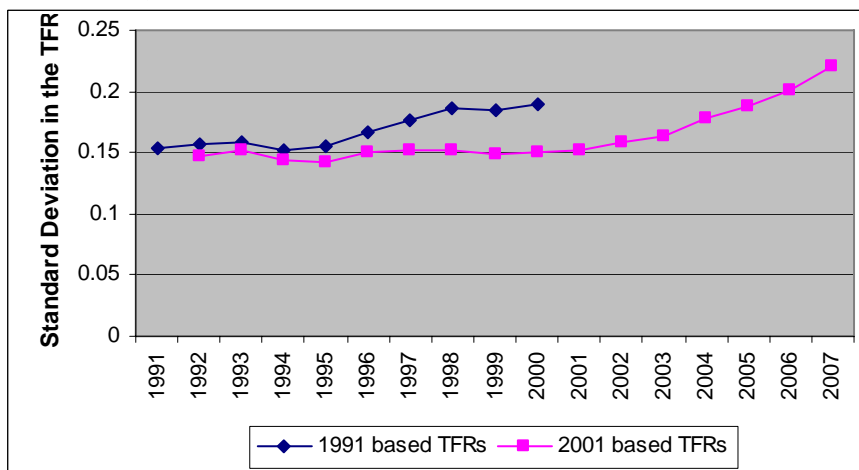
2.4 Comparing the Standard Deviation of the 1991 based and 2001 based measures

2.4.1 Fertility

The standard deviation of the 1991 based TFRs generally increases as the number of years since the census increases (figure 8).

The standard deviation of the 2001 based TFRs stays roughly constant between 1991 and 2000 (here population drift is assumed to have been corrected as a result of the 2001 based revisions). Beyond 2001 the standard deviation increases – if the increased dispersion of the local authority TFRs provides evidence of intercensal drift then this increased amount of dispersion seen between 2001 and 2007 is a concern.

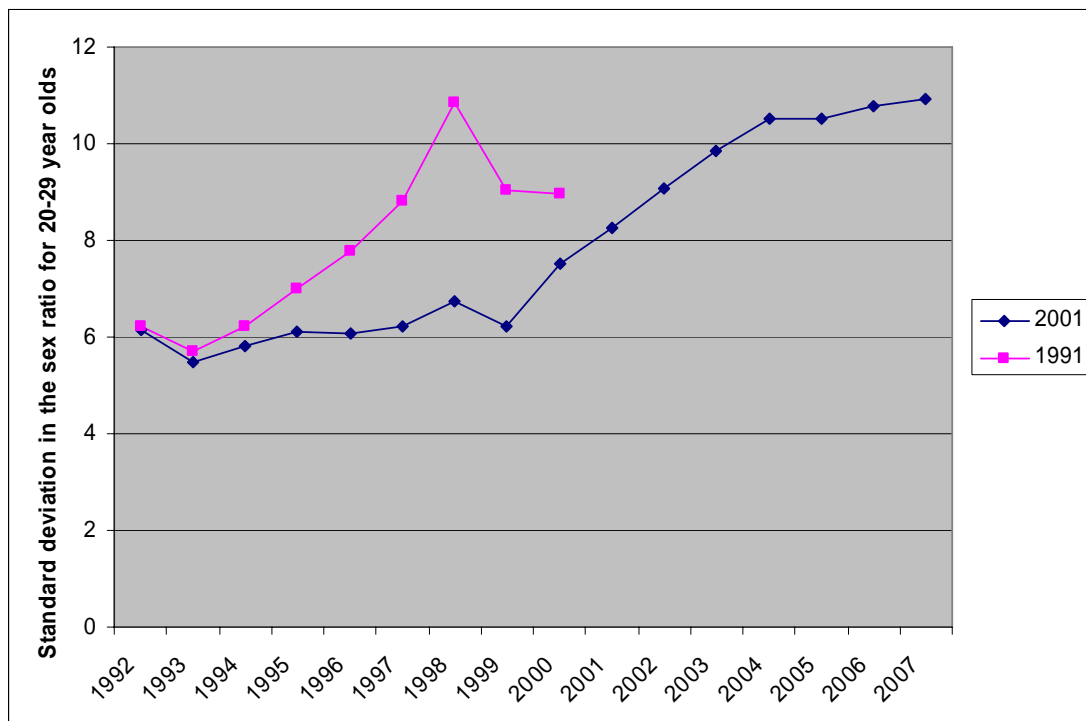
Figure 8: Standard Deviation of TFR for Local Authorities within England and Wales 1991-2007 (Isles of Scilly excluded)



2.4.2 Sex ratio

When sex ratios are calculated using the population estimates (excluding the Isles of Scilly), the standard deviation increases as the number of years since the Census increases. The standard deviation for both the 1991 and 2001 based sex ratio follows a pattern presented in Figure 9. The standard deviation's increase is greater over the period 2001-2007 for the 2001-based sex ratio as compared to the period 1992-2000 for the 1991-based sex ratio. If standard deviation and distribution of the sex ratio are a reflection of intercensal drift between censuses and the degree of drift is relative to the change in the standard deviation, then the observations for the 2001-based sex ratio pattern between 2001 and 2007 suggest a greater amount of intercensal drift in the current intercensal period, compared with the previous intercensal period.¹

Figure 9: Standard deviation in the 1991 and 2001 based sex ratio among 20-29 year olds for local authorities within England and Wales, 1992-2007(Excluding Isles of Scilly)



¹ Although as this is drift in a ratio it does not necessarily mean that numerically the drift is worse, it will depend on whether how the drift varies by sex

3. Summary of initial findings:

3.1 TFR

- An increase in the spread of all local authority TFR values may provide evidence that population drift is occurring unless another reason exists which explains the increased spread of local authority TFRs.
- A larger intercensal range in the TFR between consecutive years does tend to occur in areas where population drift identified in the 1990s was relatively high. However, this is not always the case - larger intercensal ranges can also be seen in areas where the population drift was very small.
- A larger change in the TFR between two consecutive years tends to occur in areas where the population drift in the 1990s was relatively high, however, again this is not always the case - relatively large changes in the TFR between consecutive years (between 0.2 and 0.3) are also seen to occur in areas where the population drift was very small.

Possible confounding issues include:

- Since 2001, the TFR for England and Wales has been increasing fairly rapidly. This represents a reversal of the trends observed during the 1990s. It is generally thought that when the TFR is rising there is more range for increased diversity across areas compared with when the TFR is falling and there is less scope for change.
- Changes in timing of childbearing impact upon the TFR. The timing of childbearing has changed somewhat over the period 1991-2008. These changes in timing may occur at different speed within different areas which could result in increased diversity of fertility rates across local authorities.
- Over the period 2001-2008, national fertility levels have been affected by the high levels of migration². Foreign born women, on average, have higher fertility than their UK born counterparts and migrants do not settle evenly across all local authorities. Consequently the diversity in fertility rates across local authorities could feasibly be increasing.
- Between 1991 and 2001 the standard deviation in the number of births across local authorities decreased while for the period 2001 and 2008 the standard deviation has increased (figure A2, annex A). The number of births provides an independent comparison against the TFR since births are not affected by changes to the population estimates.

3.2 Sex ratio

- the widening of the spread of the sex ratio values for 20-29 year olds during the intercensal period (1992-2000 and 2001-2007) seems to indicate the occurrence of intercensal drift
- the 2001 based sex ratio for 20-29 year olds for the period 1992-2000 gives stable variance, standard deviations and extreme values compared to the 1991 based sex ratios for the same period due to the post 2001 revisions
- the wider spread (in terms of variance, standard deviation and extreme values) of the year on year 2001 based sex ratio values for 20-29 year olds over the 2001-2007 intercensal period as compared to the 1991 based sex ratio values over the 1992-2000

² 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 pp 28-42.

intercensal period may suggest a greater level of intercensal drift for this age group in the current intercensal period.

- the range for 20-29 year olds based on the 2001 Census is narrower compared with the range based on the 1991 Census, possibly due to the shorter intercensal period

Annex

Figure A: Sex ratio pattern in mid-year estimates between 2001 and 2007

