Testing methodological changes to the household projections model

Research report

Testing methodological changes to the household projections model
Research report

If you require this publication in an alternative format please email alternativeformats@communities.gsi.gov.uk

Contents

Executive summary

The Options for the Future of the Household Projection Model report[1] provided a range of recommendations based on a combination of a methodological review, interviews and statistical testing. Communities and Local Government (CLG) distilled these into five core recommendations that are subject to further research in this report. The experiments have been carried out on the 2006-based projections so that the results can be directly compared with the 2006-based household projections published in 2009.

All of the households projections presented in this report have been produced to provide a quantitative test of the proposed new methodology. They do not form a new set of projections and the CLG 2006-based projections as published on 11 March 2009[2] remain the official household projections until their update to a 2008-base later this year.

The proposed new methodology comprises two parts. Stage 1 uses a time series projections methodology similar in many ways to that used to produce previous CLG household projections but which is based on a modified and simplified household typology. Stage 2 breaks the Stage 1 projections down into a new detailed typology using only data from the 1991 and 2001 Censuses using a method similar to that used for the Scottish, Welsh and Northern Irish projections. The proposed new typology includes detail on the number of households with children which was absent from the previous projections. The proposed new methodology does not identify concealed households.

Recommendation 1: Simplifying the projections

"That the first stage is a time series/cohort model based on standard five year age bands only with no disaggregation by household type, gender or marital status, and cohort modelling being applied at the 40-44 age band and higher".

One reason why the "Options" report recommended a simpler model with no disaggregation by household type, gender or marital status was the superior performance of projections for 2001 produced using a simple model that utilised data up to 1991. An ancillary reason was doubts about the ongoing availability of up-to-date marital status projections from ONS. Since the report was completed ONS have released 2006-based marital status projections[3] and have committed to future production so that is no longer an issue.

Rather than use a very simple age only based model, we recommend adopting a model which is disaggregated by gender and a simplified three-way relationship categorisation. The categories are couples (including married couples who are living together and cohabiting couples); separated married, divorced and widowed; and single (never married) people. This is an aggregation of the detailed categories in the existing CLG (Household Projections System, known as HOPS) model which

---

[1] http://www.communities.gov.uk/publications/housing/optionsfuturemodel

[2] http://www.communities.gov.uk/publications/corporate/statistics/2031households0309

[3] http://www.statistics.gov.uk/STATBASE/Product.asp?vlnk=14491

captures the key household formation characteristics of the relationship status groups while retaining relative simplicity.

A revised projection methodology has also been introduced which aims to project forward using more aggregate data therefore reducing the potential for errors in the underlying data to influence the resulting projections. This weights together simple and dampened logistics trends. Cohort modelling is no longer used. The simplified time-series based projections are referred to as the Stage 1 projections to distinguish them from the detailed projections by household type described below.

As with the existing HOPS system, Stage 1 projections are produced initially at the national level, then at the GOR level and finally at the LAD level with the GOR projections being controlled (constrained) to the national projections and the LAD projections being controlled to the GOR projections.

Recommendation 2: Increasing the LFS weights

"Increase the weights given to LFS data by age band in the new primary projections".

Since the "Options" paper was completed, LFS data have continued to display the sharp reductions in household representative rates for younger age groups evident since 2001 that are not fully reflected in the existing household representative rates (although there has been a partial reversal of the trend in the LFS data for 20-24 year olds since 2006). We recommend that the LFS weighting be progressively increased as the most recent LFS data approaches the 2011 Census point where it should reach a maximum of 50%. A further investigation should be instigated once the 2011 Census data become available to assess the accuracy of this method. Depending on the success of LFS data in predicting the 2011 Census point the possibility of revising the maximum weight should be assessed.

The revised projection created using the new Stage 1 methodology and the increased LFS weighting show an annual average increase in households of 247,000 between 2006 and 2031 compared with the published figure of 252,000 (which was produced using the same population base).

Recommendation 3: Census extrapolation

"That a second stage projection based on 1991 and 2001 Census data should be used to provide analysis by household type".

This recommendation has been combined with and assessed alongside Recommendation 4 given below.

Recommendation 4: Revised household types

"Revise the household type breakdown, commissioning new Census tables where required, in particular to provide information on the number of children in households".

This part of the new projections methodology is referred to as the Stage 2 projections. The results are aligned to the Stage 1 projections described above. The analysis was complicated by different definitions of household heads being used by the existing CLG (HOPS) model, and the 1991 and 2001 censuses. The recommended approach standardises the Stage 2 projections on the 2001 census definitions. Ideally, Stage 1 and Stage 2 would use identical definitions but this was not possible because of a lack of availability of time series data on a 2001 census basis and a lack of detailed household type information on the CLG definitions. At some stage in the future, possibly in the run up to the 2011 census data becoming available, we suggest that a further investigation is carried out into the possibility of "backcasting" the Stage 1 data on 2001 census definitions.

Stage 2 projects detailed headship rates (see below for the Stage 2 household typology) at the Local Authority District (LAD) level. Adjacent five-year age bands from stage 1 have been merged into 10 year age bands except for the 55 to 59 and 60 to 64 year old age bands which have been kept to provide information on pensioner households. Headship rates by age band are projected forward using a two-point exponential method.

The new Stage 2 typology has the advantage that households with children (broken down by couples and lone parents) can now be identified. The complete list of stage 2 household types is shown in table 1a on page 2.

Recommendation 5: Controlling

"Controlling sub-national projections to national projections should only be undertaken when significant new demographic data becomes available".

The current HOPS methodology as used in the 2006-based projections requires controlling to be undertaken whenever there is a change to the underlying demographic data. This is because new demographic projections change the weights of the more detailed geographies and, hence, the uncontrolled totals at a less detailed geographical area. This, in turn, will affect the final controlled results. We have tested the implications of removing the controlling procedure using different bases of both population and household representative rates. Both controlling exercises confirmed that controlling in the absence of significant new demographic data has only a marginal impact on the aggregate household projections. Note, however, that the incorporation of new LFS data still requires a variation on the controlling process and the generation of a new set of adjusted projections.

Background

In December 2008, Communities and Local Government (CLG) published a review of the methodology underlying the Household Projections System (HOPS). This review "Options for the Future of the Household Projection Model" considered the views of key experts and users and undertook some initial tests of possible methodological changes. Since this review was undertaken, CLG have published 2006-based household projections and commissioned the current project to firmly establish the quantitative impacts of the recommended methodological changes through rigorous testing on this new set of projections.

The objectives of this project were:

- To provide a quantitative assessment of the impact that the proposed methodological changes would have on the 2006-based projections.

- To recommend what changes to make to the current methodology to improve the projections.

Completion of these objectives within the specified timescales will provide CLG with further quantitative evidence on whether the new methodology should be put in place for production of the next set of household projections.

The households projections presented in this report do not form a new set of official projections and the CLG 2006-based projections as published on 11 March 2009[4] remain the official household projections until their update to a 2008-base later this year.

The current HOPS methodology has remained largely unchanged since the 1996 based household projections released in October 1999. Since then CLG have released the 2002, 2003, 2004 and 2006 based household projections.

The existing HOPS system contains a detailed breakdown which identified households represented by people who were married, single, divorced and widowed split by whether they were cohabiting or non-cohabiting (where the latter included married couple households) by 5-year age band with a further breakdown in to married couples, cohabiting couples, lone parents, single person and other multi-person households. This gives a total of 36 groups for which headship rates need to be projected (plus two concealed couple types).

While this breakdown provides a considerable level of detail there are drawbacks as identified in the "Options" project. In particular, the level of detail means that it is difficult for users to focus on the key trends and despite the level of detail some key household types, such as couple households with children were not identified.

CLG also undertook some further investigation into the needs of key users of the household projections during summer 2009. Users across central and local government and other stakeholder groups were asked to provide their views on the household types currently output by the HOPS method. 41 responses were received and whilst just over half of respondents indicated that the current household types mainly or fully met their needs, nearly three quarters wanted to see a change. There

---

[4] http://www.communities.gov.uk/publications/corporate/statistics/2031households0309

was not a strong consensus demand for any of the typologies used by the other UK countries, although the groups used in Wales proved the most popular. In summary, users would prefer size / number of children information to be made available and would like to see this in addition to the existing household types.

Tests on predictive accuracy[5] had also shown that a model with a very basic breakdown (age only) estimated using data for 1971 to 1991 would have actually produced results closer to the actual 2001 Census data than the existing HOPS system. This led to a recommendation of a two stage approach where Stage 1 would use time series (i.e. Census) data at a high level of disaggregation to project basic trends and where Stage 2 would use data from the 1991 and 2001 Censuses to provide a more detailed breakdown of household types. This forms a possible new methodology for the household projections which is being further investigated and tested in this project.

The project has also been guided by an expert Steering Group with membership drawn from central and local government, academia and private organisations. This report may not represent all views of all individual members of that group, but where possible we have responded to their valuable comments and input to the project.

---

[5] Experian (2008) – Options for the Future of the Household Projection Model: A final report. http://www.communities.gov.uk/publications/housing/optionsfuturemodel

1   Stage 1 typology

1.1   Background

The initial recommendation was for Stage 1 to have a very basic breakdown (age only). This was partly based on the predictive accuracy tests and was partly influenced by uncertainty over the future of the ONS marital status projections. Since then, ONS have produced new marital status projections and have assured us as to their future. Consequently, we have now modified the recommendation to suggest a slightly more detailed breakdown that makes use of the ONS' marital status projections and to provide a more detailed feed into Stage 2.

We now recommend the following as a simplified typology. This has the advantages of presenting a smaller and hence simpler set of groupings to aid user understanding and to minimise the potential impacts of errors in the projection data sets but which still captures the key features of household formation behaviour:

1. Couples households (here the household representative rates are defined conventionally with the male as the principle representative). This includes both married couples (where they live together) and cohabiting couples. This does not include same-sex couples

2. Female separated divorced or widowed (once married) households[6]

3. Female singles (never married) households (single as in not in a couple or separated, divorced or widowed not necessarily a one-person household[7])

4. Male separated, divorced or widowed (once married) households[6]

5. Male singles (never married) households (single as in not in a couple or separated, divorced or widowed not necessarily a one-person household).

Typically, for males, the highest household representative rates (HRR) will be for couples followed by separated, divorced or widowed, and singles but the couples HRR is effectively half of that shown as it also covers the female half of the couple whose HRR is recorded as zero under the convention used. These relativities are illustrated in Figure 1 for 35-39 year olds.

---

[6] Not necessarily a one-person household.
[7] This group, for example, will include single (not married) lone parents and people living in other multi-person households.

Figure 1: Household representative rates for 35-39 year olds
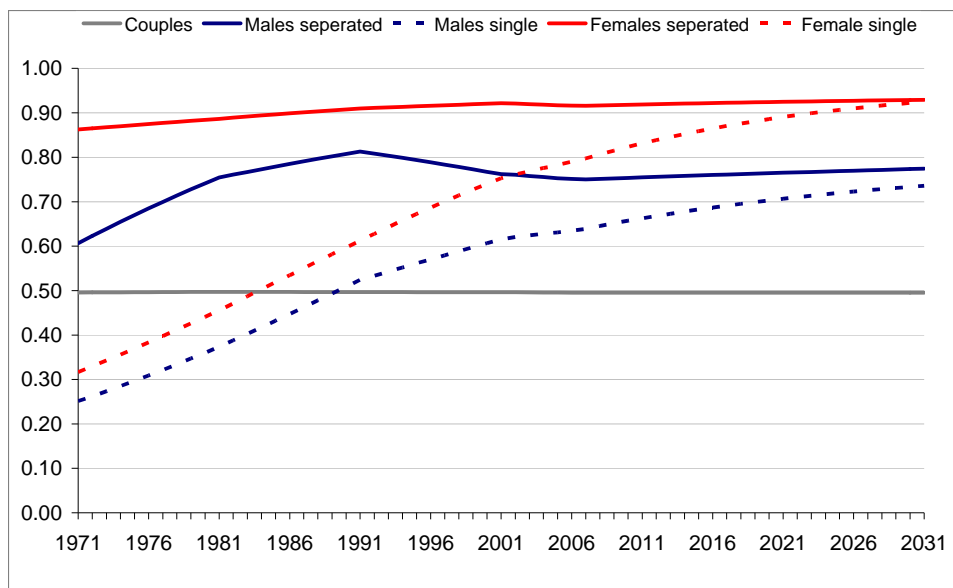
Note that changes in the aggregate HRR will be affected by relative shifts in the household population between these groups as well by the movements in the five HRRs identified. In the case of 35-39 years olds there was an increase in the aggregate HRR from 0.511 in 1981 to 0.536 in 1991 and to 0.564 in 2001 reflecting a decrease in the proportion of couples households and increases in other types, particularly singles.

In this chart, and in some subsequent charts, the convention has been adopted of showing the household representative rate for couples as being equal to the male rate divided by two. The convention used in the data used for Stage 1 is that the male is the household representative for both married and cohabiting couple households. Dividing the male rate by two emphasises that the number of housing units occupied per adult tends to be less for couples than for non-couples. Note, however, that while dividing the male rate by two is useful for illustrative purposes, it is not a strictly accurate way of showing the couple level household representative rate by age group as couples can comprise people from more than one age group and the true age-specific couple rate will actually be an average of several age groups. This discrepancy tends to be more serious for younger age groups where household representative rates for males in couples tend to be relatively low but where females may well have an older partner.

Note that the new proposed Stage 1 typology is simply an aggregation of the detailed groups in the existing CLG (HOPS) model. It uses the convention that assigns the oldest male then the oldest female if there is no male as the household representative. The main difference from the existing typology is that it combines married and cohabiting couples in to a single group. The HOPS methodology also contains a number of detailed household types such as lone parent, single person and other multi-person household which are not identified separately in the new Stage 1 typology. In HOPS, these are sub-divisions of the main household type by marital status typology and they are dealt with in the recommended Stage 2 methodology. In summary, the relationship between the new recommended typology and that used by HOPS is:

1. The recommended new methodology separates out married people who are actually living as married couples (which is one of the detailed breakdowns in

HOPS) and adds them to other cohabitees to create the new "couples" definition.

2. Adding married people who are not living as a married couple nor who are cohabiting with someone other than their spouse to the divorced and widowed groups to create the new "once married" definition.

3. Defining the new "never married" group as being equal to the old single (as in marital status not household type) non-cohabitees.

1.2   Projections

1.2.1   Household representative rates

There are two stages in generating the household projections. The initial stage projects the household population split that is needed to convert the basic projections of marital status into the three household types mentioned above.

The main issue with projecting the Stage 1 household representative rates is that there are only four observations (the 1971, 1981, 1991 and 2001 censuses) and some of those (particularly the 1991 census) look to be quite strange. We are particularly concerned about the pattern for many of the male once married but separated, widowed or divorced groups (see Appendix A) where estimated household representative rates increase sharply between 1981 and 1991 only to fall back in 2001. While we cannot say for definite that this did not happen, other household representative rates tend to evolve smoothly between Census points and it is suggestive of data errors in 1991 with the true 1991 data point being somewhere midway between the 1981 and 2001 points. We also note the very large adjustments made, as a result of under-counting, to the 1991 census estimates by ONS in generating the 1991 mid-year estimates which may be a further indicator of unreliability of some of the 1991 Census data.

We have experimented with two main ways of fitting trends to the available data:

1. A simple logistics trend - a trend fitted to $\ln (X_t / (1-X_t))$

2. A dampened logistics trends where an S-shaped curve is fitted to $\ln (X_t / (1-X_t))$[8]

It is not clear which of these is the most appropriate.

---

[8] See Appendix B

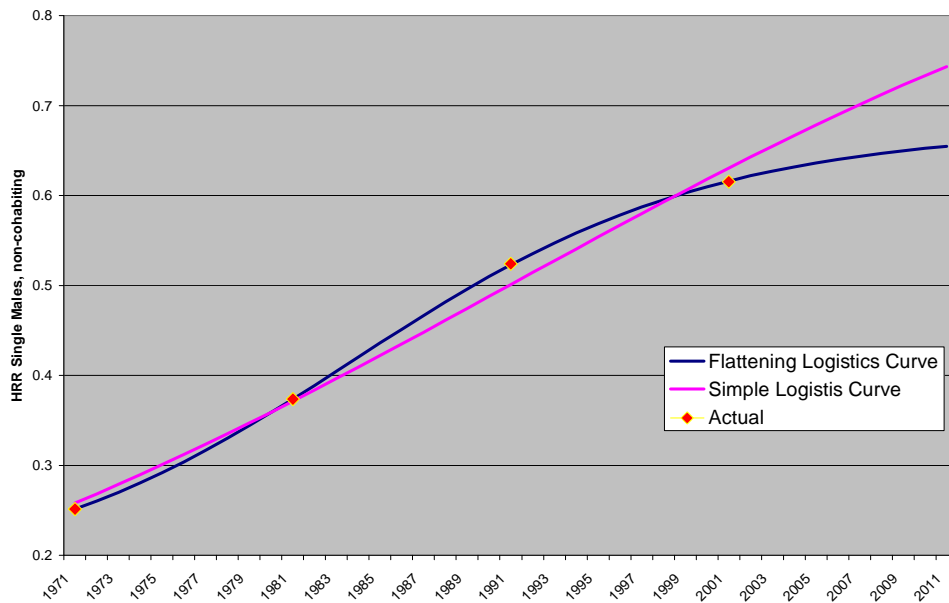**Figure 2: Single (never married) males 35 to 39 year olds**

Figure 2 shows how the estimated increase in household representative rates for this group accelerated between 1981 and 1991 but then slowed down somewhat between 1991 and 2001. This is a typical profile of many of the household representative rates not just this particular group. The dampened (referred to as "flattening" in the chart) fits this pattern very well and it implies a further flattening off of the curve after 2001. The simple logistics curve, by contrast, fails to pick up the deceleration.

Based on the data alone, we should accept the dampened logistics curve. The problem arises if we have doubts about the quality of the 1991 census data (which did have particular problems with under-recording particularly of younger single males in multi-person households which would have the effect of increasing the overall household representative rate although the household representative rates for once married but separated, widowed or divorced males and females stand out as looking particularly odd). If the 1991 household representative rate is an over-estimate then the observed deceleration between 1991 and 2001 will be exaggerated and the dampened logistics curve will incorrectly extrapolate an imaginary slowdown. In this case, the simple logistics curve may actually be a better representation of reality.

This is an intractable problem which was discussed at length in the Steering Group. Given the uncertainty, our recommendation is that we weight the alternative projections together using the following weights:

15 to 29 year olds: 80:20 weights for dampened / simple trend[9]

30 year olds and over: 60:40 for dampened/ simple trend

The reason for the differential weights is that Labour Force Survey (LFS) data indicate declining aggregate household representative rates for the younger age groups and, consequently, there is evidence that it is more appropriate to give a bigger weight to the dampened trend in these cases. This recommendation is endorsed by the Steering Group.

---

[9] The simple trend actually uses the exponentially declining weights shown in Appendix 1

The existing (HOPS) model also makes use of cohort information. There were, no doubt, sound theoretical and practical reasons for using cohort modelling in early versions of HOPS but the 2008 review found that this was only likely to be useful for groups aged 40-44 and over and that simpler a model tended to outperform HOPS in terms of predictive accuracy anyway. Given the additional concerns about the accuracy of some historical data we have not used cohort modelling for any age groups in the new recommended Stage 1 methodology.

1.2.2 LFS Adjustments

LFS data suggests that there have been some steep falls in household representative rates for some age groups since the 2001 census. For example, using smoothed LFS data the aggregate household representative rates for 15-19, 20-24, 25-29 and 30-34 year olds fell by 0.004, 0.012, 0.036 and 0.016 respectively (for charts see Appendix A). This is the equivalent of a fall of approximately 230,000 households. The current HOPS methodology does use the LFS data in projecting the household representative rates, but this is given a low weighting relative to the Census data (based on the LFS's sample size relative to that of the Census). This means that the recent and more significant falls in HRRs are not having a substantive impact on the resulting projected HRRs. If these shifts in household formation behaviour are sustained in the longer term, and this can only be truly assessed once the 2011 census results are available, the household projections would turn out to be too high using the current method.

The LFS is a sample survey and as such subject to a margin of error. However, on balance, the Steering Group felt that increased weight to the LFS data would be warranted. The LFS data has been incorporated into the England level projections:

1. The quarterly LFS household representative rate data by age (but not) sex are seasonally adjusted
2. The seasonally adjusted data are smoothed using a Henderson 9-point moving average
3. The smoothed quarterly LFS data are converted to annual series and are further smoothed using another Henderson 9-point moving average.

Adjustments are then made to all age and relationship status groups so that they move towards the smoothed LFS value with:

1. The maximum weight of 50% to reflect uncertainty over accuracy (this can be revised once we have the results from the 2011 census); and
2. The LFS weight is linked to the time since the last census (the longer the time elapsed since the last census, the less time there is for household representative rates to get back on to trend).

For example for 2007-based projections, the LFS data receives a 30% weight derived as the maximum weight (50%) multiplied by the time in years elapsed since the 2001 census divided by the maximum years between censuses (6/10). A full list of LFS weights are shown below.

Table 2: LFS weight derivation

Year	Maximum weight	Time (in years) since census	Final weight
2002	50%	1/10	5%
2003	50%	2/10	10%
2004	50%	3/10	15%
2005	50%	4/10	20%
2006	50%	5/10	25%
2007	50%	6/10	30%
2008	50%	7/10	35%
2009	50%	8/10	40%
2010	50%	9/10	45%
2011	50%	10/10	50%

Figure 3 shows the impact of the LFS adjustments in 2007. Note that the LFS adjustments have been made at the England level only and sub-national projections (produced using the weighted average of a simple and dampened trend described above) were controlled to the England totals. The largest absolute impact occurs in the 25 to 29 year old age band where the impact of the LFS sees a reduction in household representative rates of 0.011, from 0.427 to 0.415, equivalent to 2.6%.

Figure 3: Impact of the LFS adjustment on household representative rates, 2007
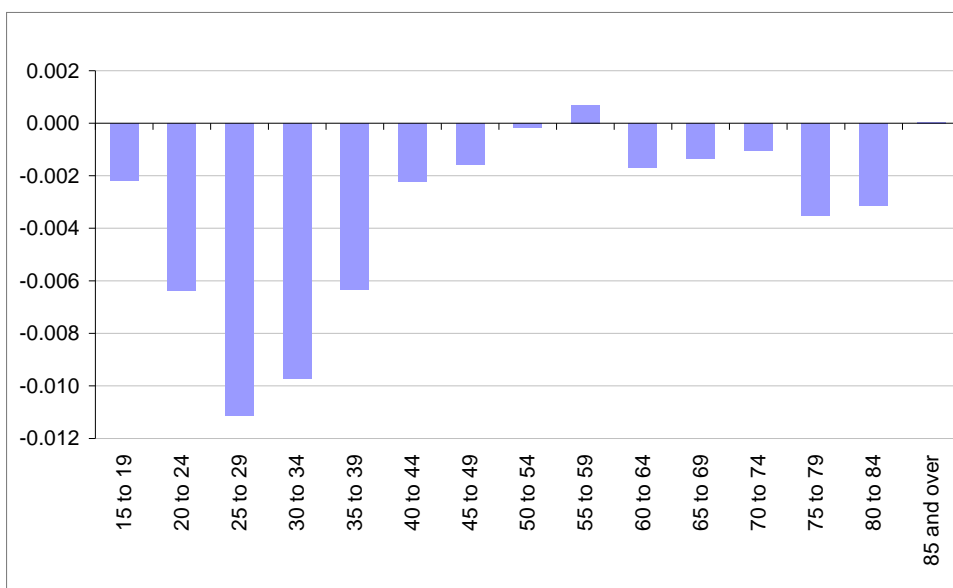
1.3   Results

Figure 4 shows the impact of the proposed changes. Overall, they produce slightly lower projections of the increase in household numbers with the new figures showing an average annual increase in household numbers of 249,000 per annum between 2007 and 2031 compared to 255,000 in the current published projections.

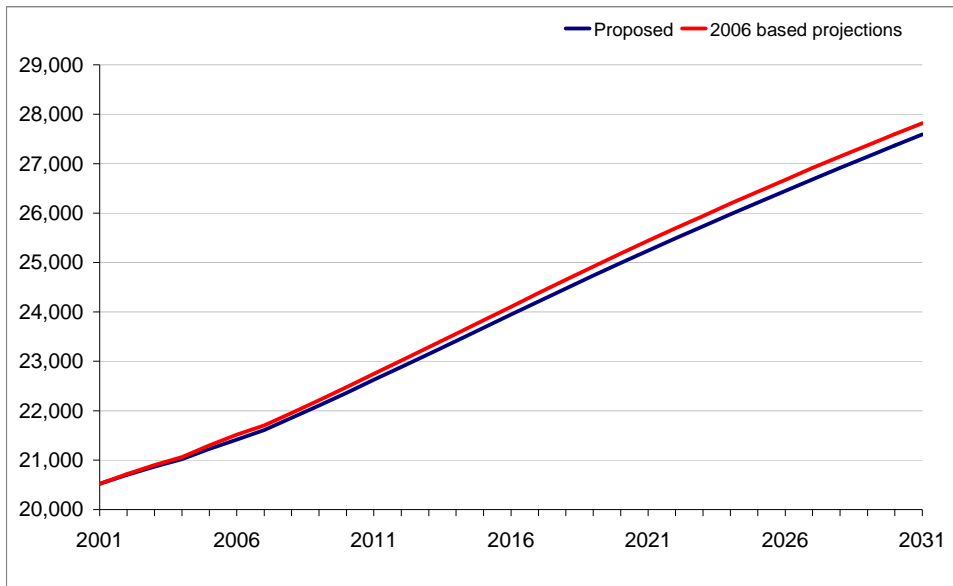**Figure 4: England projected household numbers**

Figure 5 shows the profile of the proposed adjustments. The slightly odd looking profile of the adjustments between 2002 and 2007 reflects the incorporation of the LFS adjustments for those years. The up-tick in the difference between the two sets of projections in 2007 is due to the increase in the aggregate household representative rate for 20-24 year olds in 2007 (note that the LFS adjustment has only been used for 2002-2007 in the proposed projections. LFS data are now available to 2009 and could be incorporated if users were happy with LFS adjustments that went beyond the date of the latest population estimates).

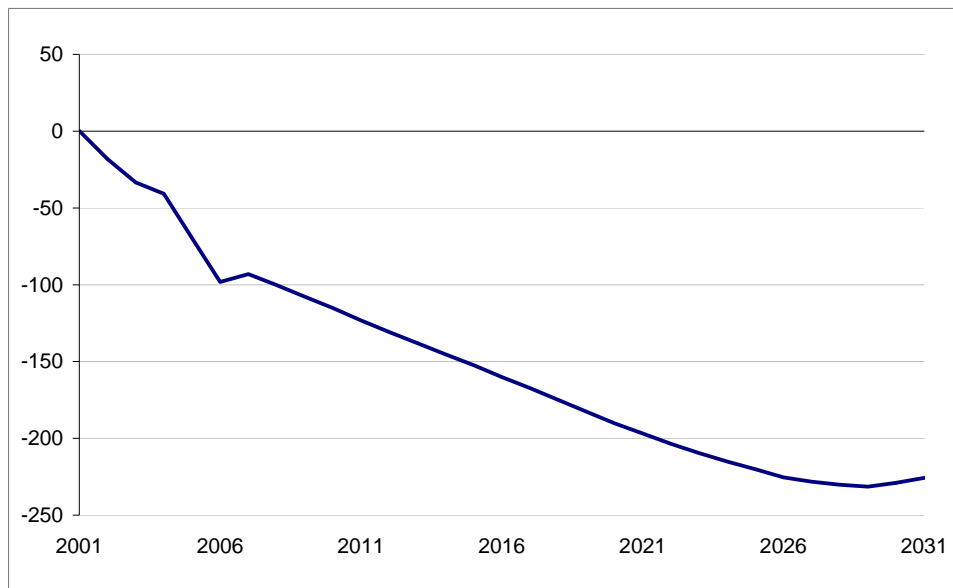Figure 5: England: Household Projections, Proposed – 2006 based

Table 3 summarises the household projections that would result using this method by government office region (GOR). The reduction in household growth at the England level is repeated in every GOR except London where the tested methodology suggests 14,000 more households (0.35%) than in the original 2006-based projections.

The increase in the projected number of households in London compared to the 2006-based projections is almost entirely due to a faster increase in female single, in particular, and, to a lesser extent, female separated, once married households.. The numbers of male single and male separated, once married households are actually less than in the 2006-based projections and the number of couples households is 5,000 higher than in the 2006-based projections by 2031. The larger number of female single households in the proposed projections is due to the new proposed methodology projecting modest increases in household representative rates for female singles in the 20-34 age band compared to some quite sharp falls in the 2006-based projections. The proposed new methodology, for example, projects a broadly stable household representative rate for single females aged 25-29 between 2001 and 3031 while the 2006-based projections have a four percentage point fall.

Table 3: Stage 1 summary results, GOR

Proposed Stage 1 Household Projections, thousands

	2001	2006	2011	2021	2031	2006-2031 p.a.
North East	1,075	1,104	1,146	1,234	1,305	8.1
North West	2,827	2,915	3,047	3,337	3,583	26.7
Yorkshire and the Humber	2,069	2,168	2,313	2,617	2,890	28.9
East Midlands	1,737	1,840	1,977	2,261	2,518	27.1
West Midlands	2,154	2,226	2,325	2,542	2,736	20.4
East	2,236	2,360	2,522	2,867	3,179	32.8
London	3,036	3,170	3,337	3,702	4,030	34.4
South East	3,294	3,432	3,610	4,017	4,389	38.3
South West	2,093	2,202	2,346	2,667	2,962	30.4
England	20,523	21,417	22,624	25,242	27,593	247.0

2006 based Household Projections, thousands

	2001	2006	2011	2021	2031	2006-2031 p.a.
North East	1,075	1,110	1,154	1,244	1,316	8.3
North West	2,827	2,931	3,067	3,367	3,617	27.4
Yorkshire and the Humber	2,069	2,181	2,332	2,649	2,932	30.0
East Midlands	1,737	1,849	1,990	2,280	2,539	27.6
West Midlands	2,154	2,237	2,340	2,566	2,762	21.0
East	2,236	2,371	2,536	2,891	3,211	33.6
London	3,036	3,178	3,337	3,697	4,016	33.5
South East	3,294	3,447	3,629	4,048	4,425	39.1
South West	2,093	2,211	2,361	2,697	3,001	31.6
England	20,522	21,515	22,748	25,439	27,818	252.1

Proposed minus 2006 based, thousands

	2001	2006	2011	2021	2031	2006-2031 p.a.
North East	0	-6	-8	-10	-11	-0.2
North West	0	-16	-20	-30	-34	-0.7
Yorkshire and the Humber	0	-13	-19	-32	-42	-1.2
East Midlands	0	-9	-13	-19	-21	-0.5
West Midlands	0	-12	-16	-24	-26	-0.6
East	0	-10	-14	-25	-32	-0.9
London	0	-8	0	5	14	0.9
South East	0	-15	-19	-31	-36	-0.8
South West	0	-10	-15	-30	-39	-1.2
England	0	-98	-123	-197	-226	-5.1

Table 4 summarises the household projections for England produced by this tested method by the revised household types. Single, never married households are expected to comprise over half (55%) of the anticipated growth in households in the projections.
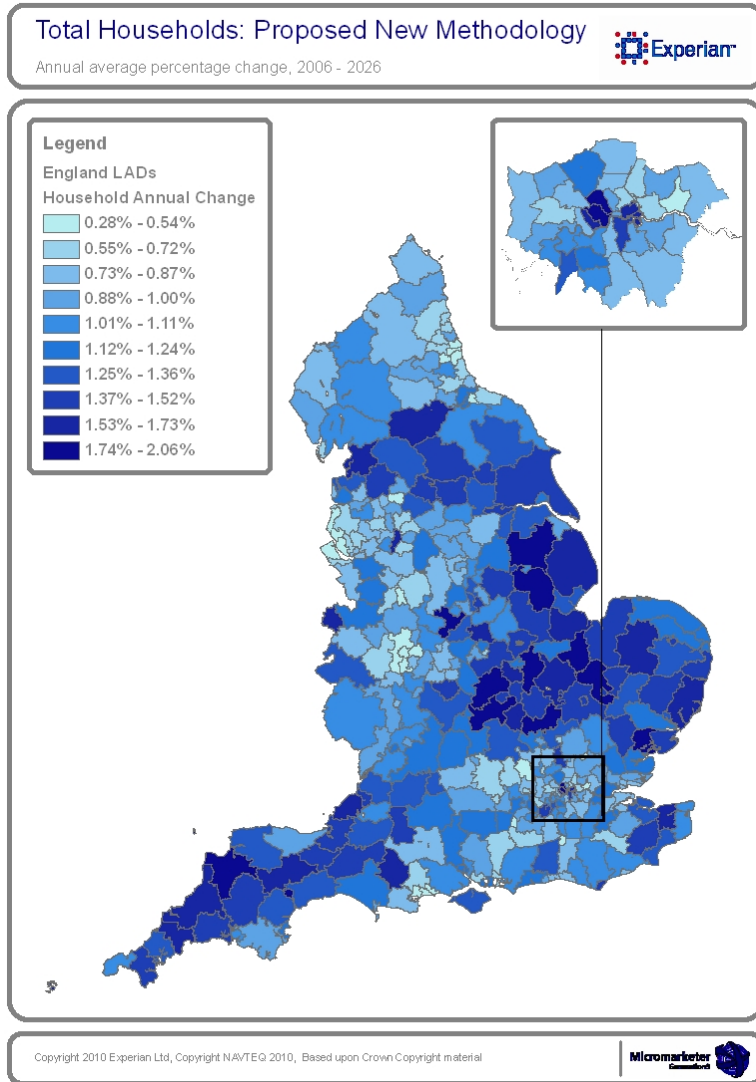
Table 4: Stage 1 summary results by type, England

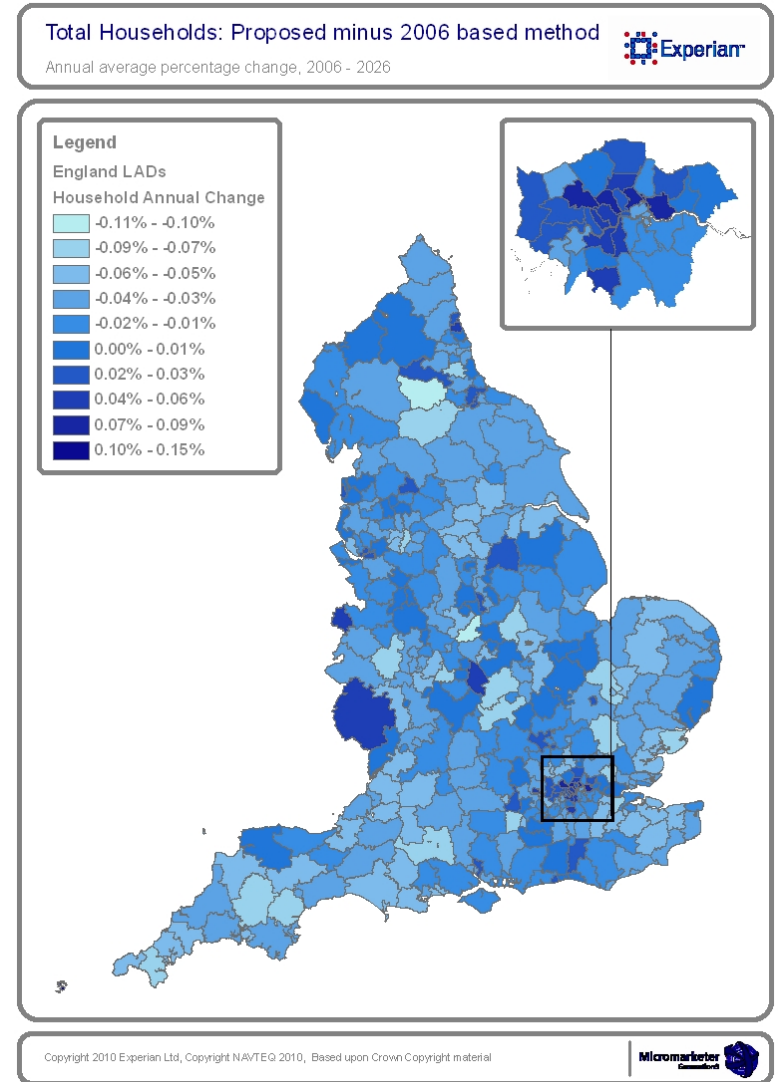| Type	2001	2006	2011	2021	2031	2006-2031 p.a.
Single, never married	3,364	3,805	4,495	5,961	7,184	135
Couples	11,497	11,551	11,792	12,416	12,985	57
Once Married, Separated or Divorced	5,661	6,060	6,338	6,865	7,423	55
Total	20,523	21,417	22,624	25,242	27,593	247

Map 1 shows that in the tested methodology, as per the 2006 based projections, it is districts within the East Midlands and the East of England (in relative terms) that are prominent in the top 10 rankings of those districts with the highest levels of household growth. The biggest increases in absolute terms are still observed in districts within Yorkshire & Humber.

The range of the change, here defined as the annual average percentage change in households from 2006 to 2026 in the tested methodology minus the 2006 based method, is small (between -0.2% and +0.2%). Map 2 shows that it is districts within Greater London where the tested method has had the greatest increase in households. There is no clear pattern for districts which have experienced a reduction in households.

Map 1

Map 2

Total Households: Proposed New Methodology

Annual average percentage change, 2006 - 2026

:Experian

Legend

England LADs

Household Annual Change

0.28% - 0.54%
0.55% - 0.72%
0.73% - 0.87%
0.88% - 1.00%
1.01% - 1.11%
1.12% - 1.24%
1.25% - 1.36%
1.37% - 1.52%
1.53% - 1.73%
1.74% - 2.06%

Copyright 2010 Experian Ltd, Copyright NAVTEQ 2010, Based upon Crown Copyright material

Micromarketer

Total Households: Proposed minus 2006 based method

Annual average percentage change, 2006 - 2026

:Experian

Legend

England LADs

Household Annual Change

-0.11% - -0.10%
-0.09% - -0.07%
-0.06% - -0.05%
-0.04% - -0.03%
-0.02% - -0.01%
0.00% - 0.01%
0.02% - 0.03%
0.04% - 0.06%
0.07% - 0.09%
0.10% - 0.15%

Copyright 2010 Experian Ltd, Copyright NAVTEQ 2010, Based upon Crown Copyright material

Micromarketer

2 Stage 2 typology

2.1 Background

Stage 2 utilises adjusted 1991 and 2001 census commissioned tables to disaggregate the household projections produced in Stage 1 into more detailed household types. This will enable the projections to provide information on size of household, particularly the number of adults and the number of dependent children in each household.

In Stage 2 we combine data with three different definitions of the household representative:

1. The existing CLG/ HOPS definition which is used in the Stage 1 projections – the oldest male then the oldest female if there is no male.

2. The 1991 Census definition – the first named person on the census form.

3. The 2001 Census definition – the eldest economically active person then the oldest inactive person if there is no economically active person.

The recommended approach for adjustments to counter the different definitions were as follows:

1. The 2001 Stage 2 household type shares are lined up with the 2001 census tables definitions.

2. The 1991 census tables are adjusted to reflect definition changes between 1991 and 2001. This has increased the number of couples households in 1991.

3. The 1991 adjustments are at the couples/ non-couples level. The share of the detailed household types within these aggregates does not change nor does the share of single person households by age.

4. The Stage 1 household totals remain the overall control total.

Ideally, Stage 1 and Stage 2 would use fully consistent definitions but this is not possible if we are to retain the link with the old time series-based model and make use of the new household typology available from recent Censuses. Nonetheless, the proposed new methodology does retain consistency with the old time series approach, and makes use of available time series back to 1971 (in Stage 1) while permitting a move to the new, and potentially more useful, household typology from recent censuses (in Stage 2).

If there is concern regarding the use of different definitions in Stage 1 and Stage 2 then we recommend that the bulk of any published detailed material be on the Stage 2 (i.e. 2001 Census consistent) definition with Stage 1 only being available on request. Note that this does not affect total household numbers which are consistent in Stage 1 and Stage 2. Publication of projections by household type which are consistent with the published 2001 Census table may actually help to alleviate any confusion already caused by the current inconsistency between the household types in the existing CLG (HOPS) projections and the publically available Census data.

2.2 Methodology

Stage 2 therefore utilises adjusted 1991 and 2001 census commissioned tables to disaggregate the household projections produced in stage 1 into more detailed household types. This will enable the projections to provide information on size of household, particularly the number of adults and the number of dependent children in each household.

Stage 2 initially works with data at the Local Authority District (LAD) level. Adjacent five year age bands from stage 1 have been merged into 10 year age bands (except for the 55 to 59 and 60 to 64 year old age bands which have been kept to provide information on pensioner households). Stage 1 outputs are used to constrain the LAD figures.

The proportions of households by household type and age group of the head of household are derived from the adjusted census tables for 1991 and 2001. This data is supplemented by data on non household reference persons by age band. The proportions of each household type and non household reference persons, known as the headship and non headship rates, sum to one within each age band.

The headship and non-headship rates by age band are projected forward using a two-point exponential method, replicating the method used by GRO Scotland[10]. The same technique has been used to provide headship and non headship rates for 1992 to 2000. The headship and non-headship rates are constrained so that they cannot individually go above 1 or below 0 and that they sum to 1 within each age band.

The first cut of Stage 2 LAD level household projections are calculated by applying headship rates to the household population projections by age band to give an estimate of the number of heads of household for each household type and age band. The total number of households is adjusted to be consistent with stage 1 household projections at the broad age band and LAD level.

2.3 Minimum adults check

The first cut household projections are tested to ensure the minimum number of adults required to fill the projected households is not greater than the projected adult private household population from the population projections minus an allowance for the institutional population. This was not found to be an issue at the LAD level and subsequent regional and national levels. No further adjustment is made for any period. This test is also run after the dependent children adjustment is made, again no problems were identified.

---

[10] Household Projections for Scotland 2006-based, GRO Scotland, October 2008

2.4   Dependent children adjustment

A second check is run to ensure the minimum number of dependent children as suggested by the projected household types does not exceed projections of dependent children in the population. This check required an estimate and projection of the number of dependent children for 2002 to 2031. By definition, dependent children includes all 0 to 15 year olds so we have used single year ONS Sub National Population Projections for this element. However, 16-18 year olds require a further assumption concerning the proportion of the 16 to 18 total population that are dependent children (i.e. single and inactive and a student). This has been calculated at the LAD level by pushing forward 2001 census shares with growth rates in the number of full time students in the corresponding age band. Full time student data has been sourced from the LFS as a proxy to capture changes in the levels of dependent children due to increasing participation in post 16 education. Proportions have been kept fixed past the last data point (2009).

The test was initially run at the regional level where only in Yorkshire and the Humber was the minimum number of dependent children found to be greater than projections of dependent children in the population. Yorkshire and the Humber is unique in having experienced growth in the minimum number of dependent children as suggested by the household estimates in the 1991 and 2001 censuses combined with falling estimates of dependent children post 2001. Extrapolating the growth in households with dependent children forward, combined with falling dependent children projections is causing the minimum number of dependent children projections to be greater than dependent children projection from 2007 onwards. The other eight regions complied with the test.

In the stage 2 projections, the test was run at the local authority level and within each district an adjustment was made to the number of households with children so that the ratio of the actual number of dependent children to the implied household projection outcome of dependent children is constant in the future. In the Bradford example shown below, we have adjusted the number of households with children down to achieve the constant ratio. In the Darlington chart the adjustment to households with children is upwards to ensure the constant ratio. The divergence between the minimum dependent children projection and the dependent children projection in the Darlington example, whilst feasible, would require a significant increase in the proportion of households with more than three dependent children for the household projections to fit with the population projections. In each case the number of households without children is also adjusted to maintain overall consistency with the stage 1 household projections.

In Scotland, for the revised 2006-based household projections, adjustments were required for Aberdeen City and Edinburgh. Adjustments are made in Scotland when the tests fail, whereas we have adjusted each LAD in England to maintain a constant ratio, at 2001 levels, between the actual number of dependent children to the implied household projection outcome of dependent children.

Figure 6: Dependent children adjustment example, Bradford
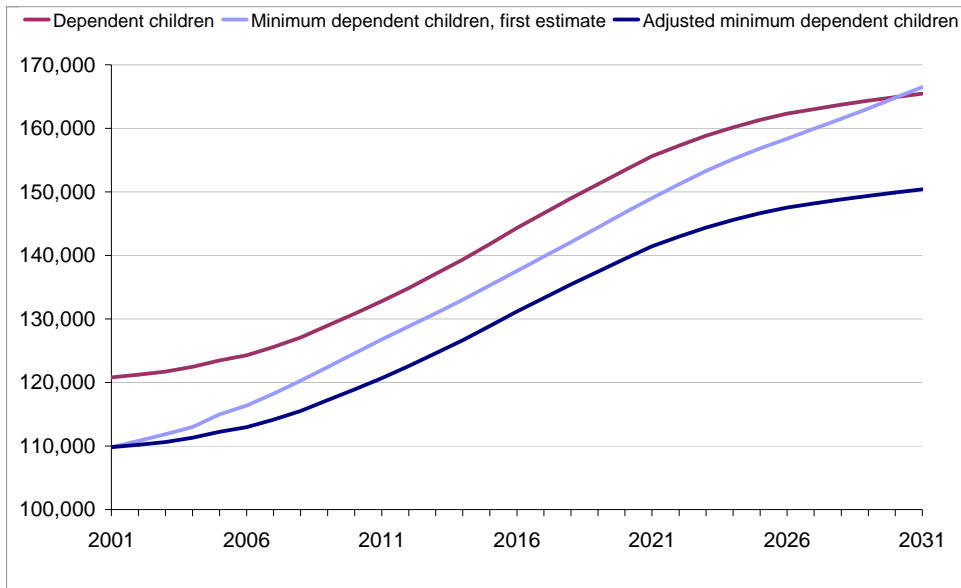
Figure 7: Dependent children adjustment example, Darlington
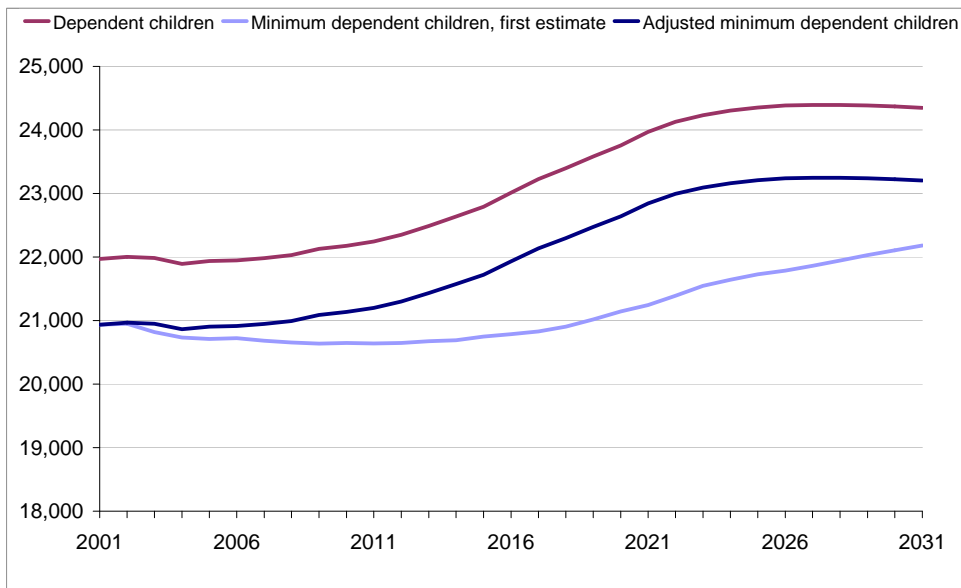
2.5   Results

Stage 2 results indicate that household growth in England will be focused within both male and female one person households and to a lesser extent within one family and no other couple households with no dependent children. This is unchanged from the first cut household projections, i.e. those produced before the dependent children adjustment.

Table 5: Stage 2 results, England

Type		2001	2006	2011	2031	2006-2031 p.a.
One person households	Male	2,678	3,113	3,565	5,279	87
	Female	3,626	3,937	4,292	5,826	76
One family and no others	Couple: No dependent children	5,434	5,764	6,153	7,488	69
	Couple: 1 dependent child	1,265	1,229	1,242	1,271	2
	Couple: 2 dependent children	1,702	1,611	1,539	1,415	-8
	Couple: 3+ dependent children	750	730	712	727	0
	Lone parent: 1 dependent child	592	687	803	1,206	21
	Lone parent: 2 dependent children	398	442	494	705	11
	Lone parent: 3+ dependent children	185	204	229	334	5
A couple and one or more other adults	No dependent children	1,532	1,404	1,297	978	-17
	1 dependent child	460	399	368	296	-4
	2 dependent children	200	189	186	185	0
	3+ dependent children	98	100	105	129	1
Lone parent and one or more other adults	1 dependent child	162	172	188	244	3
	2 dependent children	67	76	86	129	2
	3+ dependent children	34	40	47	75	1
Other households		1,341	1,321	1,319	1,305	-1
Total		20,523	21,417	22,624	27,593	247

The projected increase in lone-parent households is greater than in the 2006-based projections largely because of a bigger increase in household representative rates for younger single (and, to some extent, once-married) females in the new Stage 1 projections. In these cases, the 2004 and 2006-based household representative rate projections more closely resemble the Stage 1 dampened trend projections rather than the average of the dampened and simple trends finally used in the new projections. The levelling off in the 2004 and 2006-based projections is due to cohort effects (i.e. they extrapolate an earlier levelling off in the household representative rates for younger age groups between 1991 and 2001). Cohort effects were not included in the new Stage 1 process largely because the Options for the Future of the Household Projection Model report found that they did not add to predictive accuracy, particularly for younger age groups. The new Stage 1 projections show a continued increase, although at a declining rate, in the household representative rates for the younger single (never married) female groups beyond 2001 and this translates in to more single female households in Stage 1 and, consequently, more lone parent households in Stage 2. Given the importance of the link between Stage 1 single female households and Stage 2 lone parent households, it may be worthwhile experimenting with the weights further in future but in order to do this it would be useful to have an external reference point, not related to the household projections, to base the experiments on.

One of the advantages of the proposed methodology is the ability to aggregate household types to examine for example, households with dependent children. This is shown in Map 3, districts within the Greater London and a corridor of districts north of London in the East of England and the South East are expected to experience the highest growth rates.

Map 3

Households with dependent children
Annual average percentage change, 2006 - 2026

Experian

Legend

England LADs

Household Annual Change

- -0.53% - -0.14%
- -0.13% - 0.09%
- 0.10% - 0.27%
- 0.28% - 0.43%
- 0.44% - 0.59%
- 0.60% - 0.75%
- 0.76% - 0.94%
- 0.95% - 1.22%
- 1.23% - 1.64%
- 1.65% - 2.35%

Copyright 2010 Experian Ltd, Copyright NAVTEQ 2010, Based upon Crown Copyright material

Micromarketer
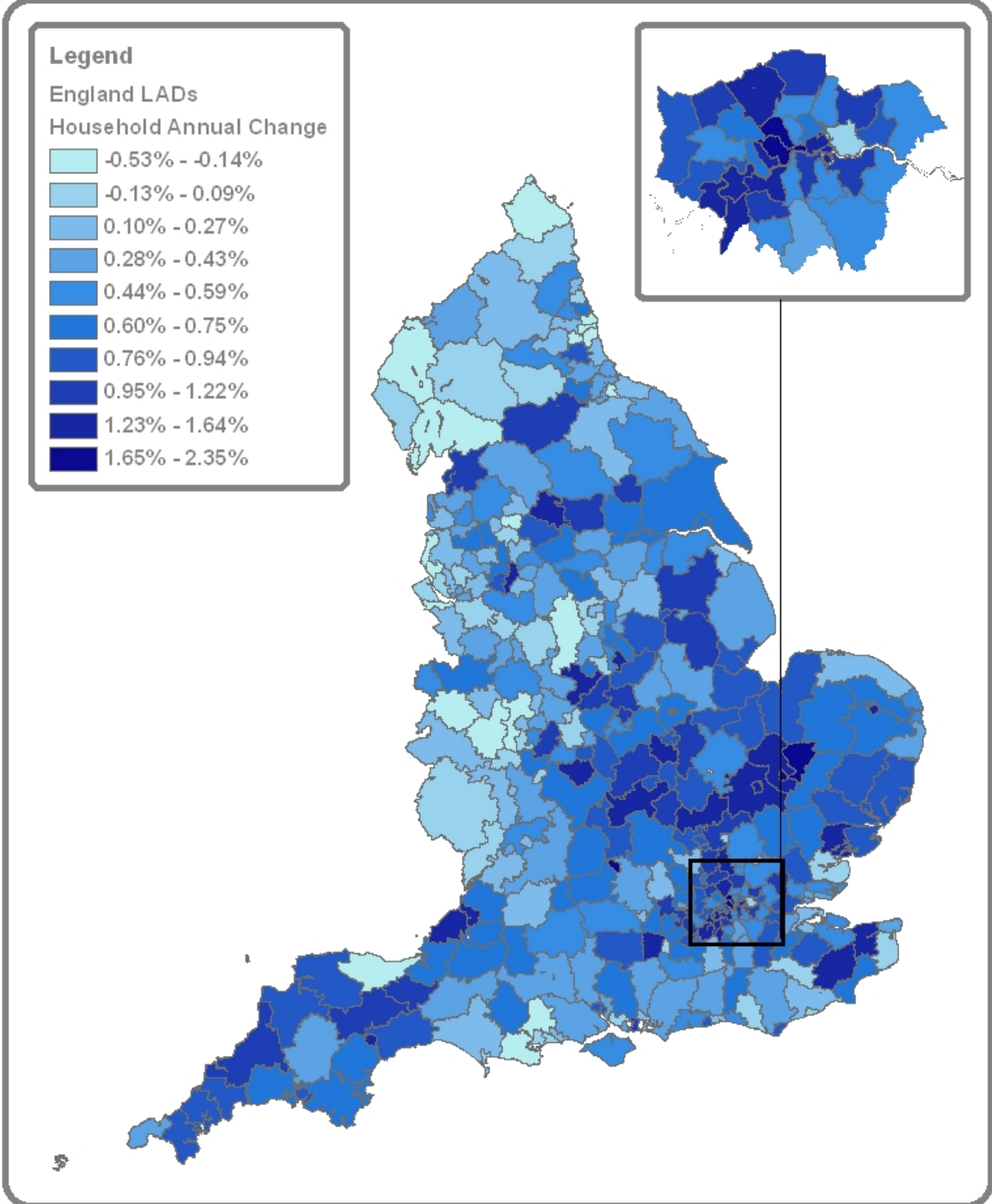
3   Controlling

3.1   Background

The 2006-based methodology requires controlling to be undertaken whenever there is a change to the underlying demographic data (population or another input into household representative rates). By changing the controlling procedure it would be possible to run the household projections as soon as the new sub-national population projections become available, removing one potential source of a time lag between population and household projection data release. The "Options" report therefore recommended that controlling sub-national projections to national projections should only be undertaken when significant new demographic data becomes available. In this chapter we test the implications of removing the controlling procedure.

3.2   Controlling – Test 1

Test 1 uses 2006 household population projections and the 2004 based household representative rates for England, the regions and local authorities within Yorkshire & Humber. In this test we assess the impact of applying 2004 versions of the regional HRRs to the 2006 population projections without constraining to the national totals and compare the summation of the regional household projections against the household projections produced using national HRRs.

Table 6: Controlling test 1, households, England 2026 (000s)

	Sum of regions	National HRRs	Difference
15_19	134	134	0
20_24	781	780	2
25_29	1,554	1,550	4
30_34	2,313	2,309	4
35_39	2,703	2,702	0
40_44	2,499	2,500	-1
45_49	2,118	2,118	0
50_54	2,100	2,100	0
55_59	2,304	2,304	0
60_64	2,270	2,270	0
65_69	1,953	1,953	0
70_74	1,702	1,703	0
75_79	1,732	1,732	0
80_84	1,287	1,287	0
85&	1,365	1,365	0
Married couple	9,135	9,132	3
Cohabiting couple	3,576	3,576	0
Lone parent	2,037	2,039	-2
One person	10,245	10,233	12
Other multi-person	1,822	1,828	-5
Total	26,816	26,808	8

The household projection in 2026 produced using the sum of the regional HRRs was found to be only 0.03% higher than the equivalent national HRR produced series. Table 6 shows that this is equivalent to an additional 8,000 households. The differences are concentrated in 25 to 29 and 30 to 34 age bands and one person household types.

The test was also run at the district and regional geographic level with Yorkshire and Humber used as the example region. A similar set of results were found; the household projections as a result of summing the districts was found of be marginally higher (0.06%) than what would be the case if we used a top down approach and used the regional HRRs.

3.3   Controlling – Test 2

The second test reverses the inputs from test 1 in that it uses 2006 based HRRs combined with 2004 based household population. It assesses the impact of using current HRRs constrained with an old vintage of population data.

Table 7: Controlling test 2, households, England 2026 (000s)

	Sum of regions	National HRRs	Difference
15_19	115	115	0
20_24	735	737	-2
25_29	1,425	1,429	-4
30_34	2,112	2,115	-3
35_39	2,497	2,497	0
40_44	2,326	2,326	1
45_49	2,074	2,074	0
50_54	2,079	2,078	0
55_59	2,297	2,297	0
60_64	2,276	2,276	0
65_69	1,959	1,959	0
70_74	1,714	1,714	0
75_79	1,727	1,727	0
80_84	1,267	1,267	0
85&	1,236	1,237	0
Married couple	8,892	8,894	-3
Cohabiting couple	3,422	3,423	-1
Lone parent	1,872	1,870	2
One person	9,885	9,896	-11
Other multi-person	1,770	1,765	5
Total	25,840	25,848	-8

In the second test the household projection in 2026 produced using the sum of the regional HRRs was found to be 0.03% lower than the equivalent national HRR produced series, equivalent to a difference of 8,000 households. The differences are again concentrated in the young adult age bands and one person household types.

Both controlling exercises confirm that the controlling procedure should only be undertaken when significant new demographic data, such as sub national population projections become available. When the next round of national and sub-national population projections are released we recommend controlling procedures are implemented. In producing the 2006 projections, the systems were designed to run the controlling procedures relatively quickly making the driving factor to remove them largely redundant. The tests confirm that when only a small aspect of the household projection methodology is changed, such as the inclusion of latest LFS data, the impact of the controlling procedure is small.

4 Testing the sensitivity of the projections to changes in marital status projections

To test the sensitivity of the household projections to changes in marital status projections, we have re-run the 2006-based household projections, incorporating the latest (2006-based) ONS marital status projections. The 2006-based marital status projections were not available at the time the 2006-based household projections were produced, so 2003-based projections were used, although these were adjusted to rectify an increase in the discrepancy between the number of married males and married females arising from applying the revised sex and age distribution from the 2006-based population projections to an earlier vintage of marital status projections.[11]

ONS 2006-based marital status projections were published 31 March 2009 for the period 2007 to 2031 and are produced for England & Wales together. The projections are produced for total population for the following categories:

- never married

- married

- widowed

- divorced

The HOPS model was re-run for England and for each of government office regions (GORs). Whilst marital status projections are produced at a national level (England & Wales) only, the regional sensitivity to marital status projections is likely to vary given the existing regional marital status structure.

The changes between the 2006 and 2003 based marital status projections are largely based in the younger age groups due to underestimation of migration in the 2003 based projections, elsewhere the results were reasonably comparable.

4.1 Total adult population method

The first stage of the testing involves deriving total population by marital status as follows:

---

[11] The marital status balancing procedure kept the aggregate gap between the number of married males and females at the same level as in the revised 2004-based projections. This affected both the marital status shares and the household population/institutional population relativities, but it did not have a big impact on the final results. Household levels in 2031, without the marital status adjustment, were 23,000 households higher (equivalent to an increase of 0.08%) than the final national level projections.

1   Source unit level legal marital status and cohabitation projections from ONS, 2007 to 2031 (the projections are published rounded to the nearest hundred or thousand depending on the variable).

2   The 2006 marital status projections do not include figures for 2006. There is also no published estimate of cohabitation for 2006. Given this, for the base year (2006), published marital status estimates have been used for 2006 marital status projections. No similar estimates exist for cohabitation, so the base year has been back cast using the 2007 to 2031 projections.

3   As stated earlier the marital status projections are only available for England and Wales combined. The England & Wales growth rates were therefore used to push forward the marital status profile for England.

4   The next stage involves constraining the government office regions to the new marital status profile for England.

Note that throughout this process, total population remains unchanged from the 2006 based population projections, only the marital status profile has changed.

4.2   Institutional population

The institutional population is derived accordingly:

1. The England under 75 institutional population remains unchanged. It is fixed at 2001 census levels by sex & marital status & age cohort as per the methodology used in the 2006 household projections.

2. The England over 75 age bands institutional population changes with the changing marital status profile.

3. The final step constrains the GOR institutional population to the new total for England.

4.3   Household population and households

The new total population and institutional population are then used to derive household population consistent with the 2006 marital status projections. Household representative rates are then applied to the new household population to produce new household projections.

1. Household population is equal to total population minus institutional population at sex and marital status and age level.

2. England households are derived using the same household representative rates (HRRs) as used in the 2006 based household projections;

3. Adjustments are made to GOR level HRRs for internal consistency and then applied to GOR level household population.

4.4   Results

Figure 8 below shows the results of including the latest marital status projections in the 2006 household projections methodology. Incorporating the new marital status projections results in a total of 4,400 additional households by 2031 when compared with the 2006 household projections using the 2003 marital status projections. This represents a difference of just 0.02% to the published total in 2031.

Figure 8: Total households, 2006 household projections and new marital status projections
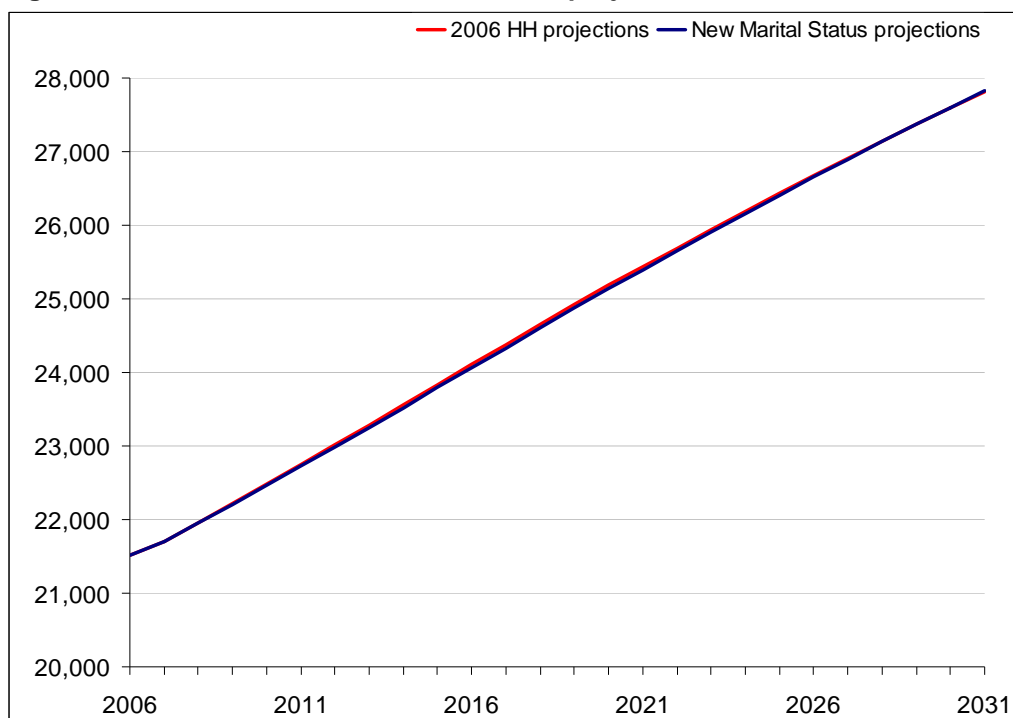
4.4.1   Results by marital status

Table 8 presents total households by marital status for the new marital status projections and the published 2006 household projections in 2031. Incorporating the latest marital status projections has a marked difference on the number of households by marital status, with an increase in married and divorced households, offset by a decline in the number of widowed households and to a lesser extent, single households. The differences are due to the changes in the marital status population projections and how these impact on household representative rates by age and gender.

Table 8: Households by marital status (thousands)

Marital Status	Base 2006	New Marital Status projections 2031	2006 household projections 2031	Difference (2031)
Married Total	10,407	10,423	10,296	127
Single Total	5,509	10,673	10,747	-74
Widowed Total	2,623	2,388	2,671	-283
Divorced Total	2,975	4,339	4,105	234
TOTAL	21,515	27,823	27,818	4

4.4.2 Results by government office region

The marital status projections are only available at a combined England and Wales level, but the changes between the 2006 and 2003 marital status projections does impact upon the household projections at regional level, due to the existing marital status in each region. Table 9 shows that the biggest increase in terms of both level and percentage change terms is in Greater London with 6,000 additional households relative to the published 2006 household projections, followed by Yorkshire and the Humber, with 2,000 additional households. Elsewhere the differences are small, with slight declines in the number of households in the South East, East of England and the East Midlands.

Table 9: Total households by region

Region	Base 2006	New Marital Status projections 2031	2006 household projections 2031	Difference (2031)
North East	1,110	1,316	1,316	0
North West	2,931	3,618	3,617	1
Yorkshire and the Humber	2,181	2,934	2,932	2
East Midlands	1,849	2,538	2,539	-1
West Midlands	2,237	2,761	2,762	0
East	2,371	3,210	3,211	-1
London	3,178	4,022	4,016	6
South East	3,447	4,423	4,425	-1
South West	2,211	3,001	3,001	0
England	21,515	27,823	27,818	4

The impact of incorporating the 2006 based marital status projections at the aggregate level (the total across all marital statuses) has had only a negligible impact on the household projections. The impact is subdued as the comparison is with the 2006 based household projections, which adjusted the 2003 based marital status projections to the changing age and gender profile prevalent in the 2006 based population projections. It is therefore not a direct comparison with 2003 based marital status projections. There were however differences identified in this exercise in the household projections by marital status. Future marital status projections should always be incorporated into the household projection methodology as long as marital status remains a component of the methodology.

Appendix A

Household representative rate charts

Stage 1 household representative rates: 15 to 19 year olds
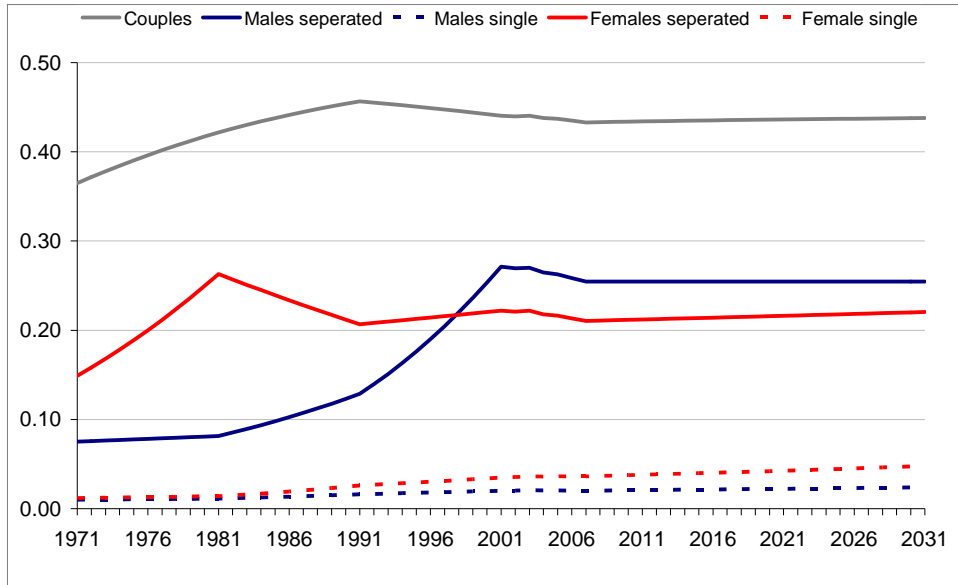
Stage 1 household representative rates: 20 to 24 year olds
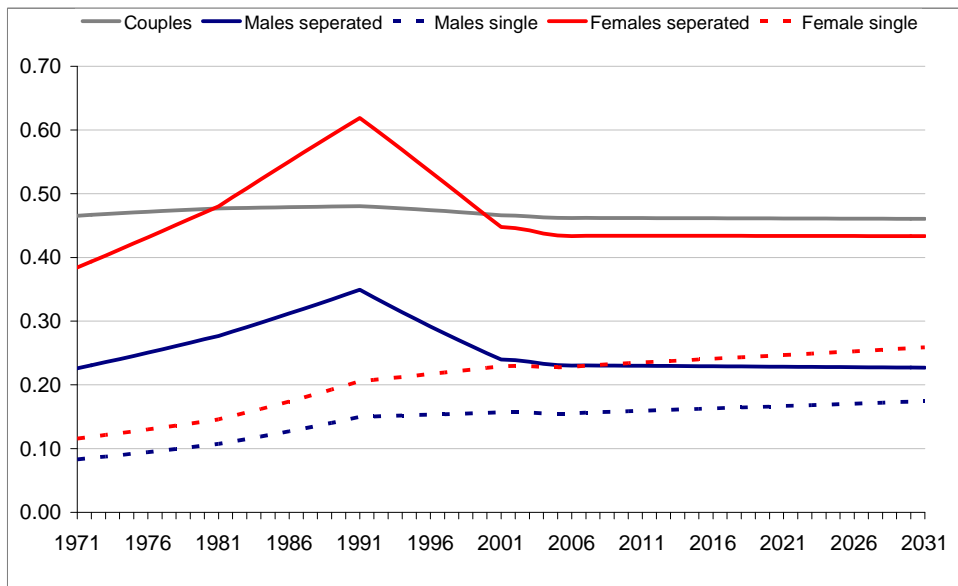
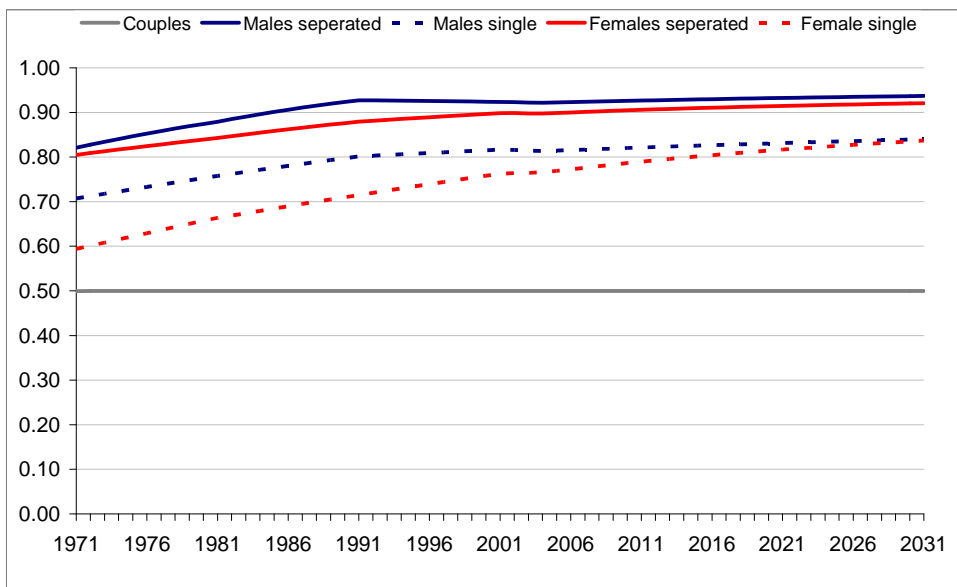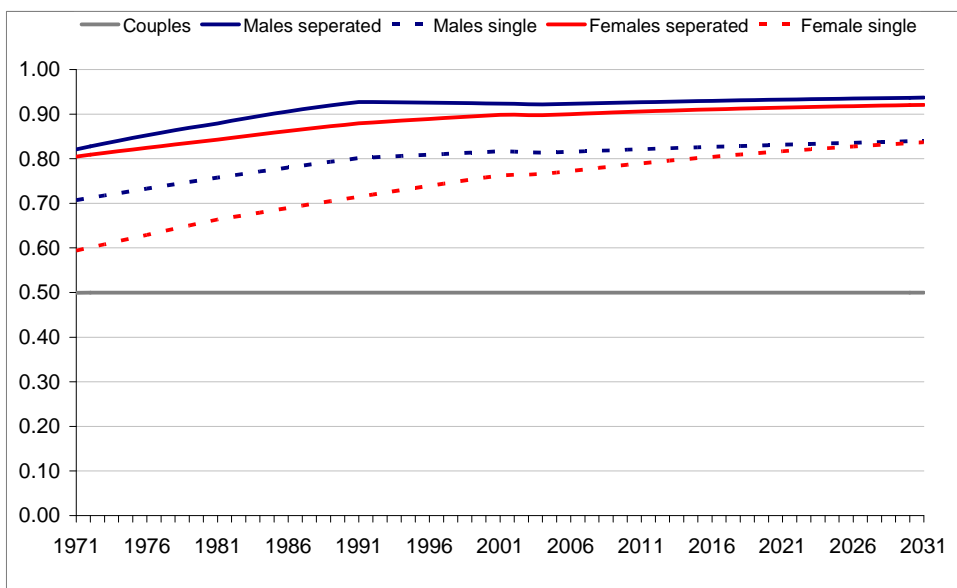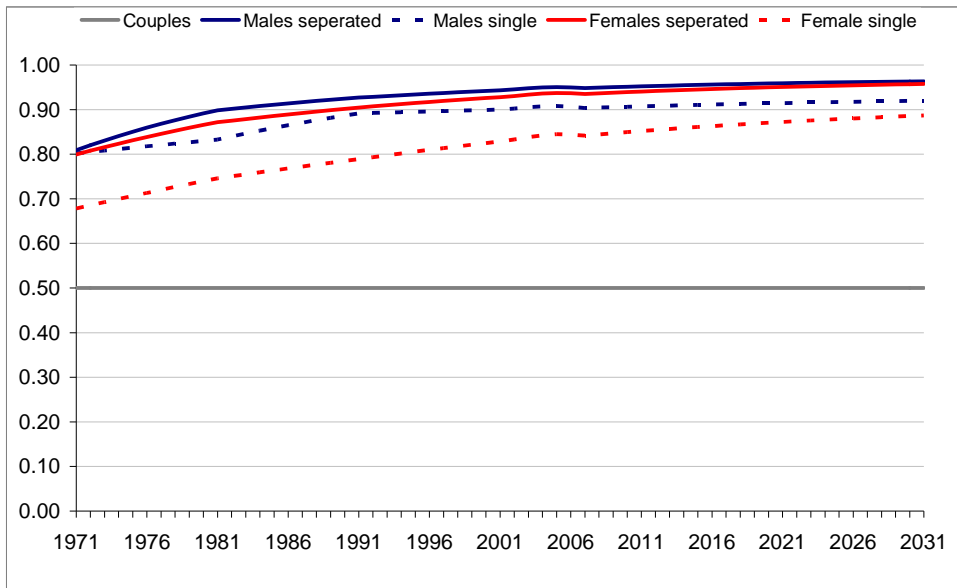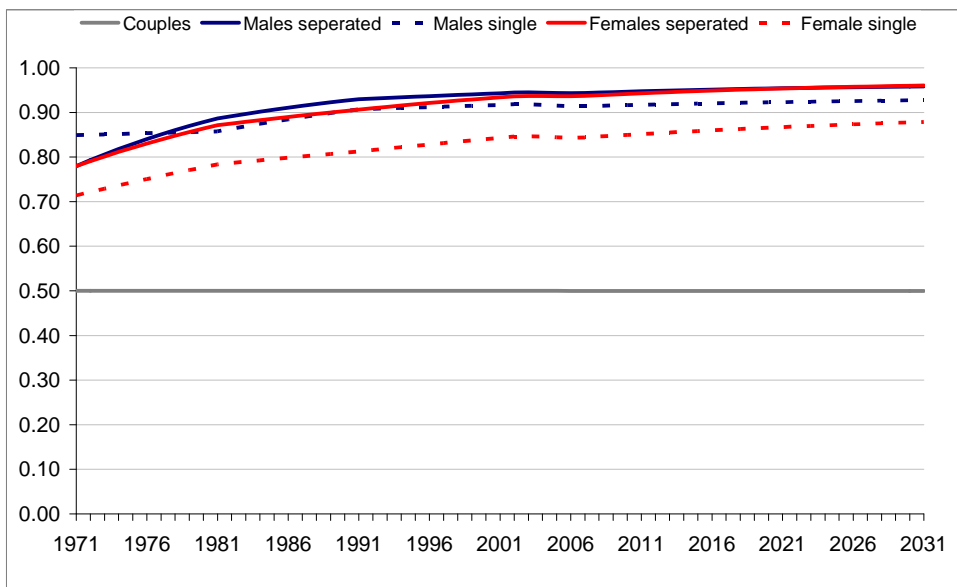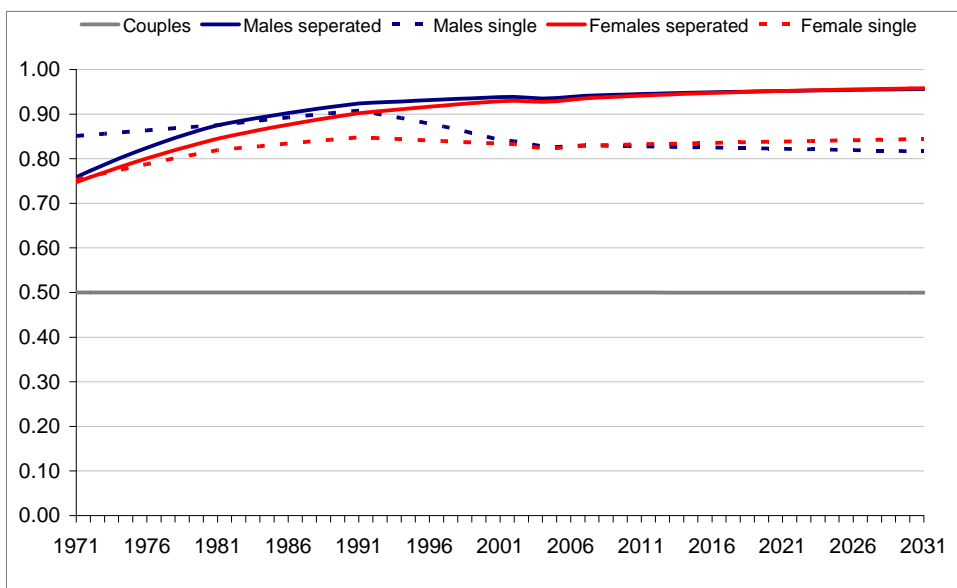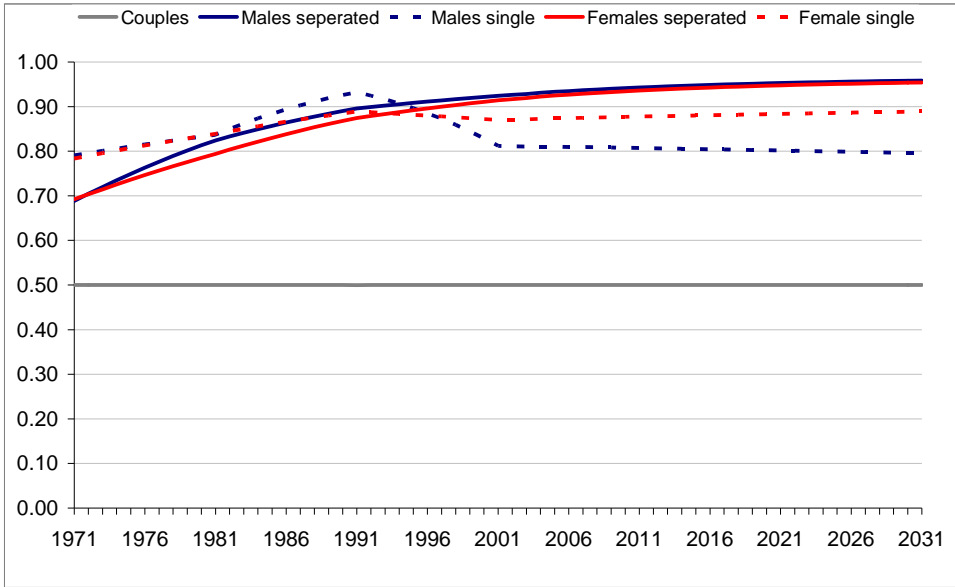## Stage 1 household representative rates: 25 to 29 year olds
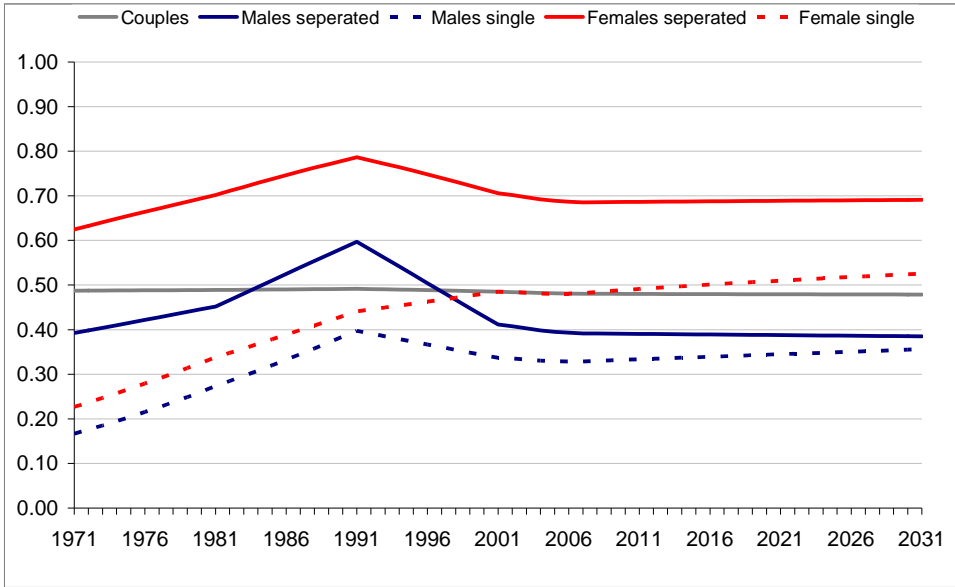
Stage 1 household representative rates: 30 to 34 year olds
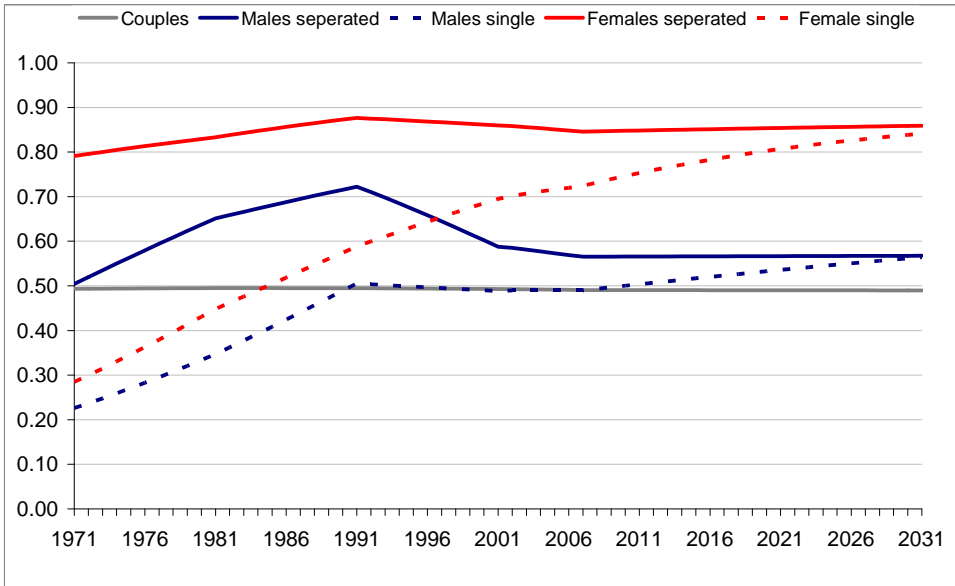
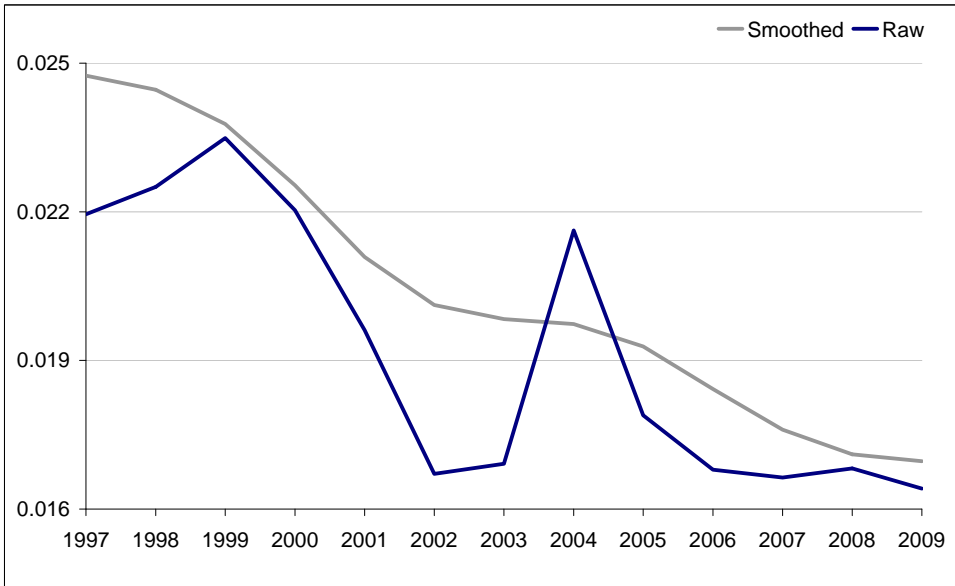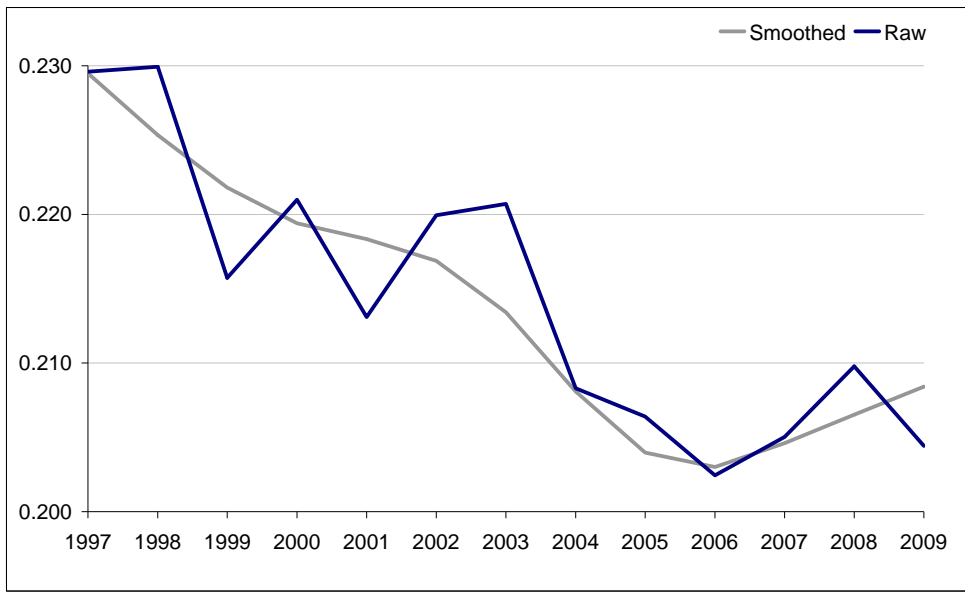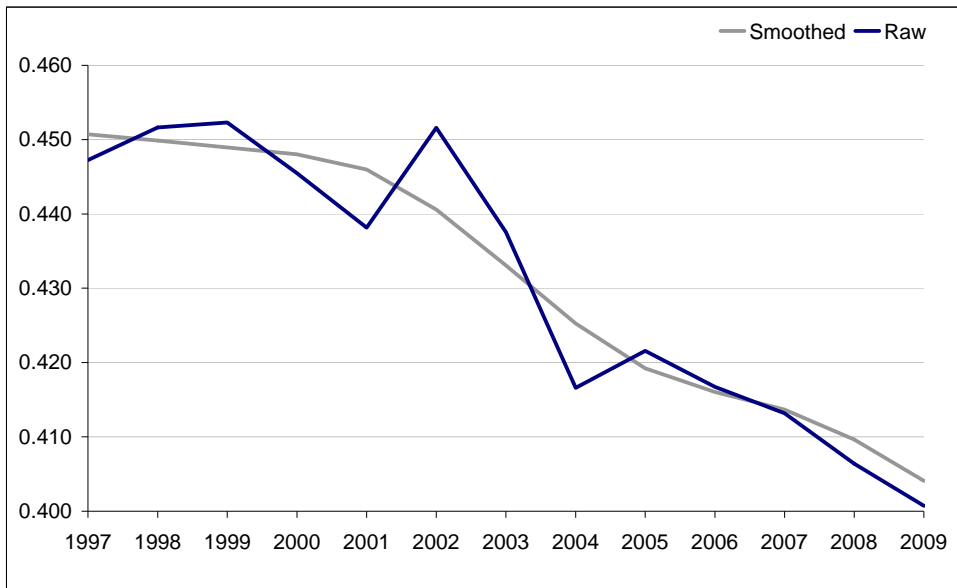## Stage 1 household representative rates: 35 to 39 year olds
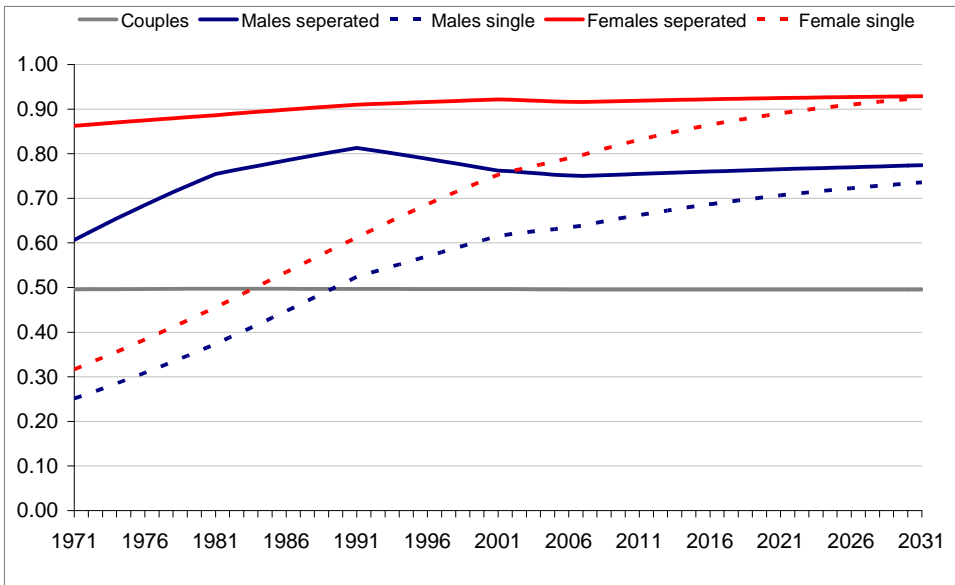
Stage 1 household representative rates: 40 to 44 year olds
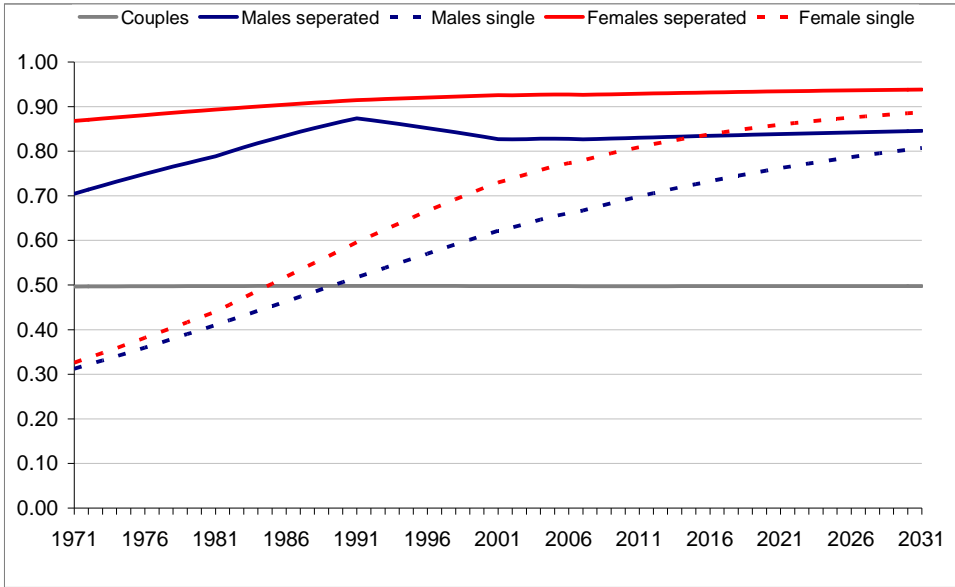
Stage 1 household representative rates: 45 to 49 year olds
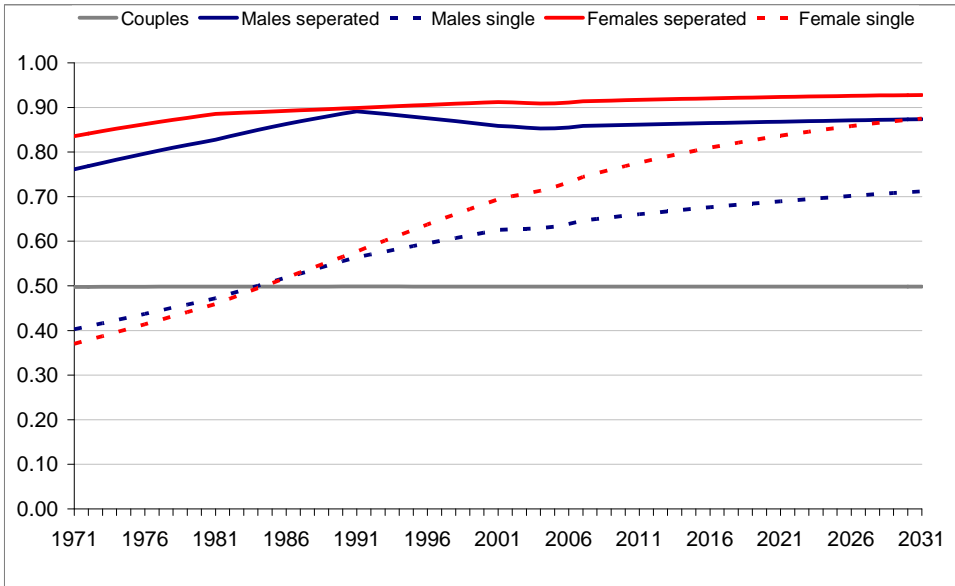
Stage 1 household representative rates: 50 to 54 year olds
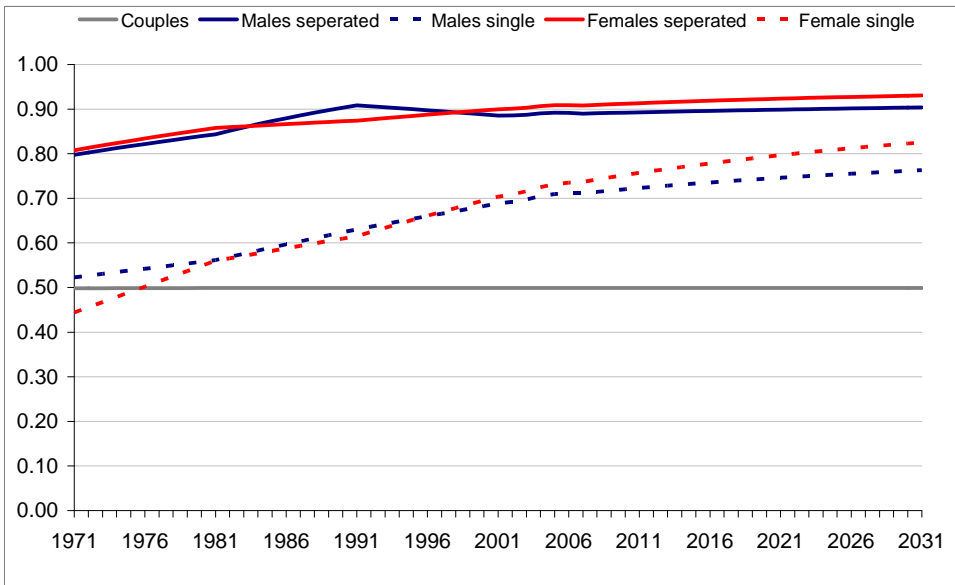
Stage 1 household representative rates: 55 to 59 year olds
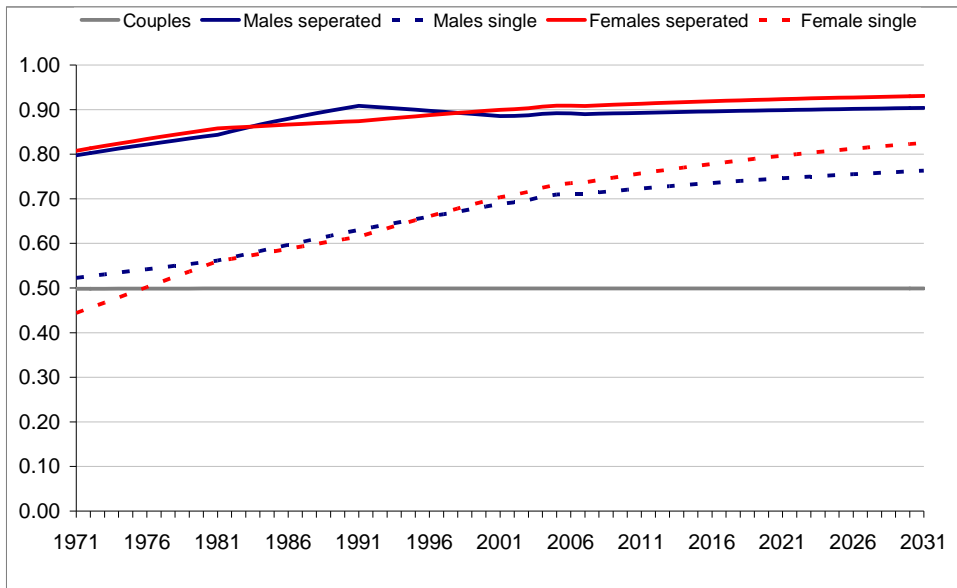
Stage 1 household representative rates: 60 to 64 year olds

Stage 1 household representative rates: 65 to 69 year olds

Stage 1 household representative rates: 70 to 74 year olds

Stage 1 household representative rates: 75 to 79 year olds

Stage 1 household representative rates: 80 to 84 year olds

Stage 1 household representative rates: 85& year olds

LFS household representative rates: 15 to 19 year olds

LFS household representative rates: 20 to 24 year olds

LFS household representative rates: 25 to 29 year olds

LFS household representative rates: 30 to 34 year olds
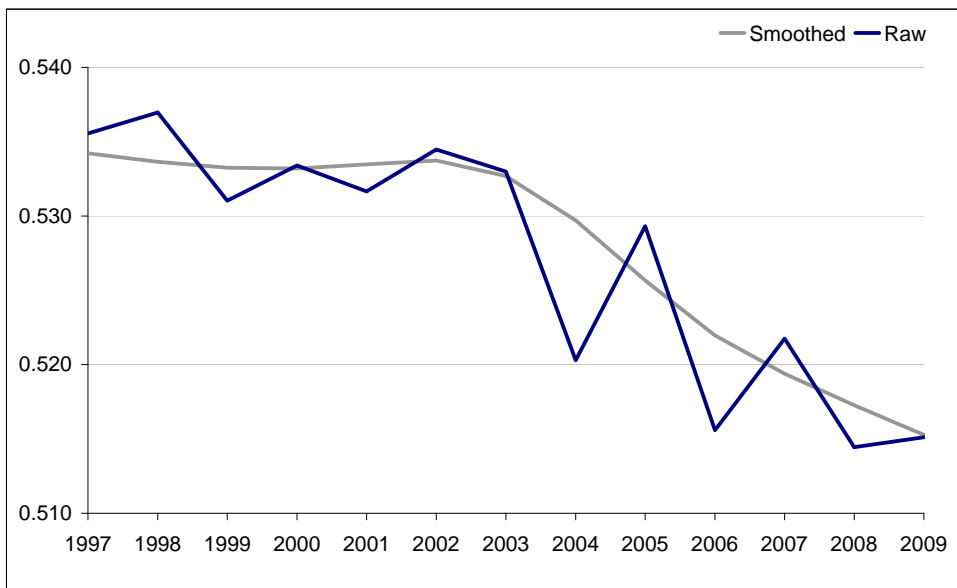
Appendix B

Trend lines

The **dampened logistics trend** is defined as:

$$\ln(X_t / (1 - X_t)) = \beta_o + \beta_1 / (1 + \exp(-1.\beta_2(t - \beta_3)) + e_t$$

Where X_t is the ratio of interest, beta 1..4 are estimated coefficients and e_t is the equations error term.

The **simple logistics** extrapolation is defined as:

$$\Delta \ln(X_t / (1 - X_t)) = 0.42(\ln(X_{2001} / (1 - X_{2001})) - \ln(X_{1991} / (1 - X_{1991})))$$
$$+ 0.33(\ln(X_{1991} / (1 - X_{1991})) - \ln(X_{1981} / (1 - X_{1981})))$$
$$+ 0.25(\ln(X_{1981} / (1 - X_{1981})) - \ln(X_{1971} / (1 - X_{1971})))$$

Appendix C

Additional stage 1 maps

Single Households
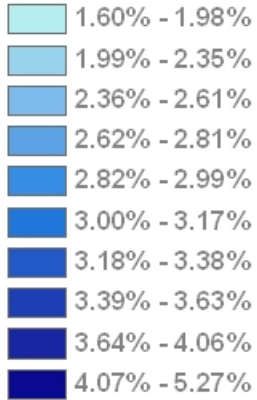
Annual average percentage change, 2006 - 2026

Legend

England LADs

Household Annual Change

- 1.60% - 1.98%
- 1.99% - 2.35%
- 2.36% - 2.61%
- 2.62% - 2.81%
- 2.82% - 2.99%
- 3.00% - 3.17%
- 3.18% - 3.38%
- 3.39% - 3.63%
- 3.64% - 4.06%
- 4.07% - 5.27%

Micromarketer

Couple Households

Annual average percentage change, 2006 - 2026

Experian™

Legend

England LADs

Household Annual Change

- -0.49% - -0.29%
- -0.28% - 0.00%
- 0.01% - 0.17%
- 0.18% - 0.31%
- 0.32% - 0.45%
- 0.46% - 0.59%
- 0.60% - 0.75%
- 0.76% - 0.92%
- 0.93% - 1.14%
- 1.15% - 1.58%

Micromarketer
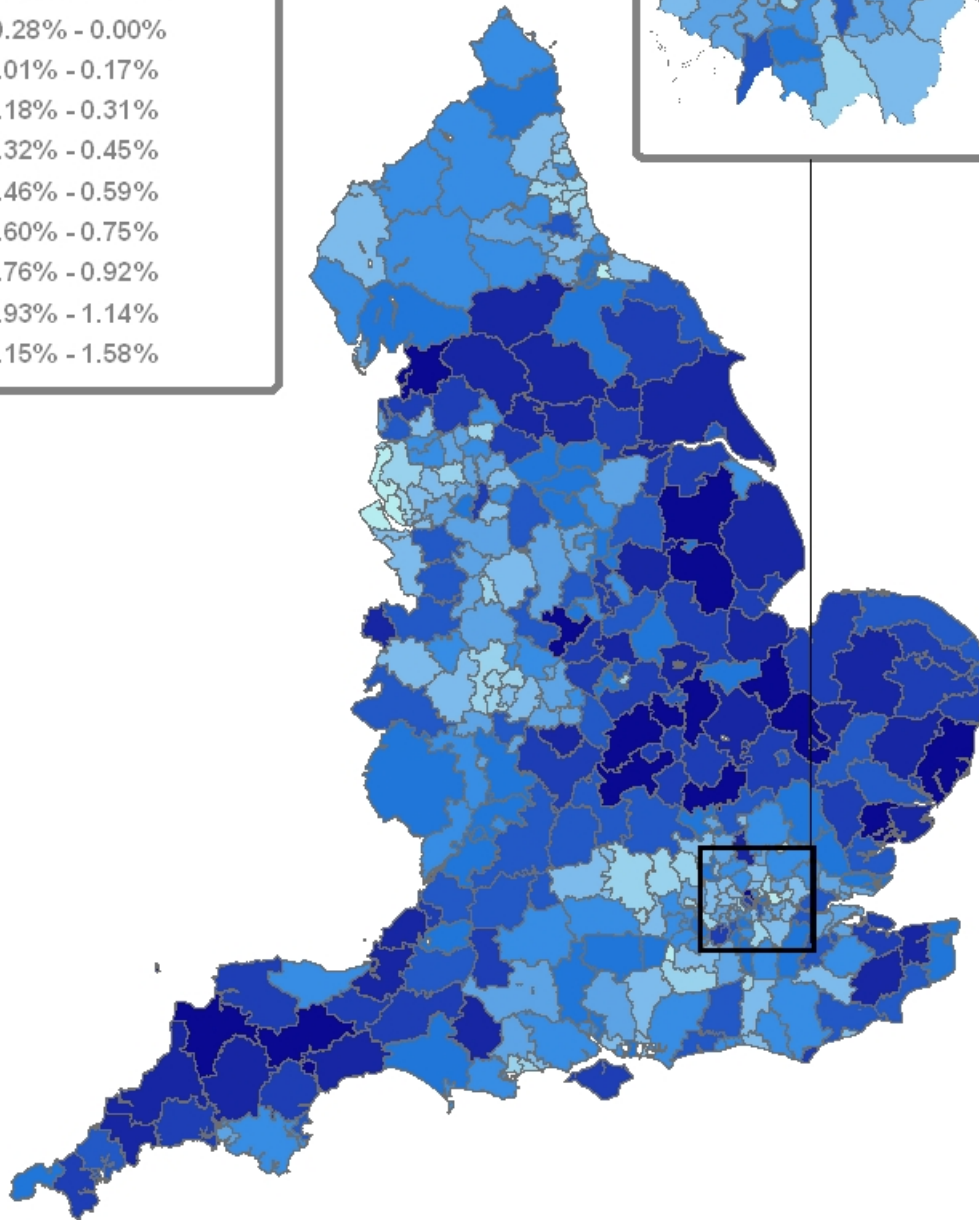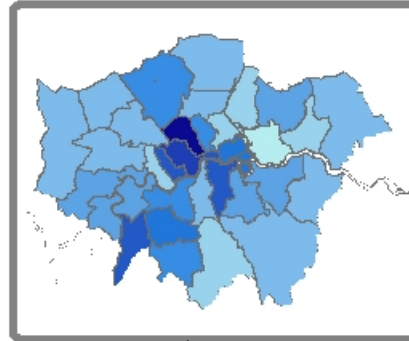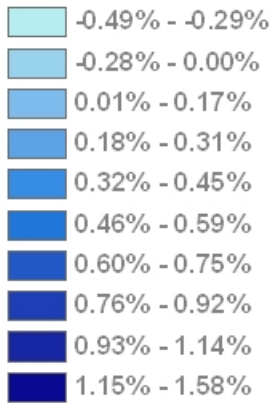
Once-Married Households

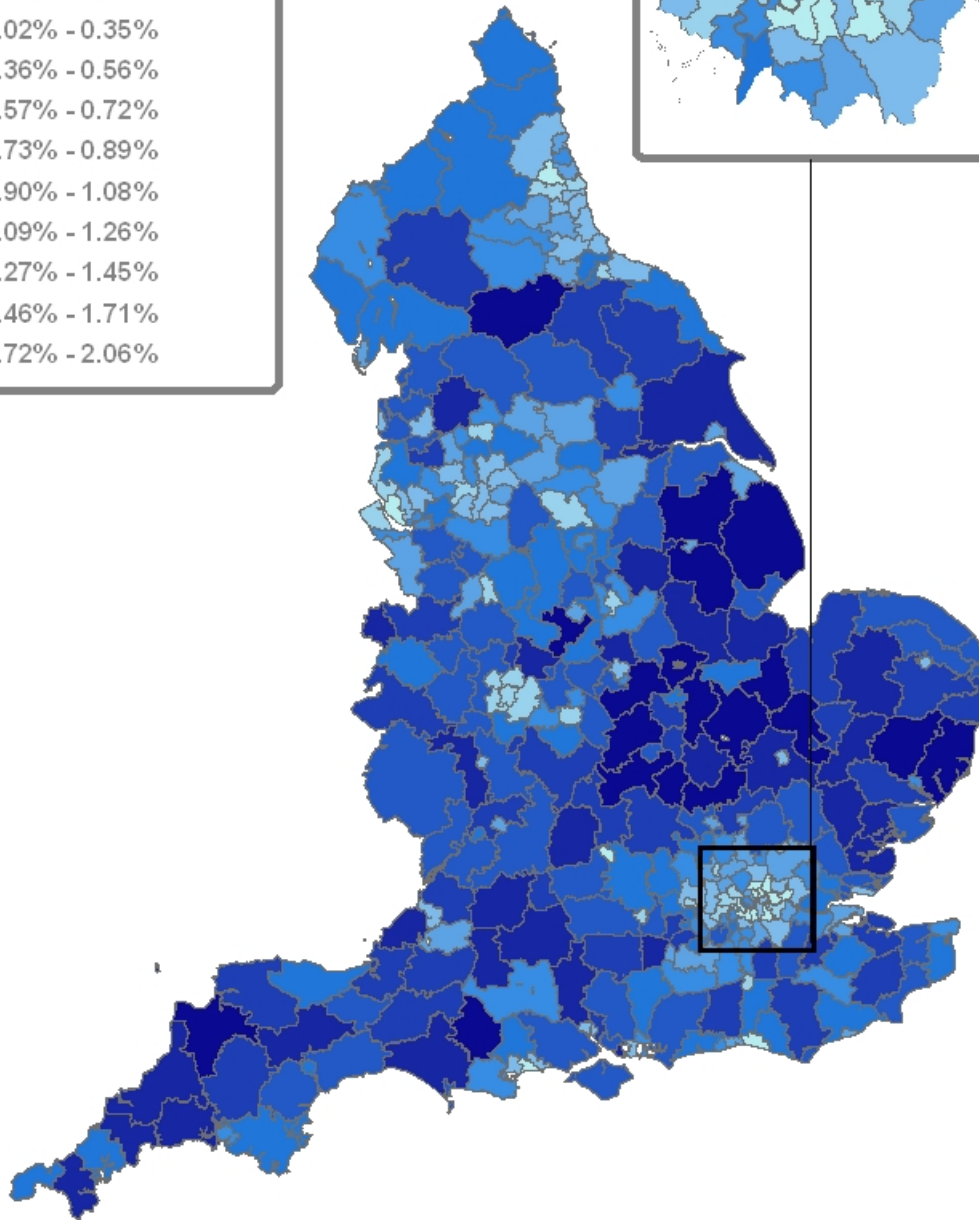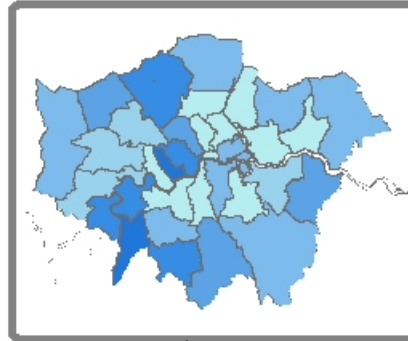Annual average percentage change, 2006 - 2026

Experian™

Legend

England LADs

Household Annual Change

- -0.42% - 0.01%
- 0.02% - 0.35%
- 0.36% - 0.56%
- 0.57% - 0.72%
- 0.73% - 0.89%
- 0.90% - 1.08%
- 1.09% - 1.26%
- 1.27% - 1.45%
- 1.46% - 1.71%
- 1.72% - 2.06%

Micromarketer

Appendix D
Bibliography

Options for the Future of the Household Projection Model: A final report, CLG, December 2008
http://www.communities.gov.uk/publications/housing/optionsfuturemodel

Household Projections to 2031, England, CLG, March 2009
http://www.communities.gov.uk/publications/corporate/statistics/2031households0309

Marital Status projections, ONS, March 2009
http://www.statistics.gov.uk/STATBASE/Product.asp?vlnk=14491

Household Projections for Scotland 2006-based, GRO Scotland, October 2008
http://www.gro-scotland.gov.uk/files1/stats/gros-household-projections-for-scotland-2006-based/gros-household-projections-for-scotland-2006-based.pdf