The value of census data for population projection

Phil Rees and Pia Wohland

Census 2011: impact and potential
Exploring the research potential of the 2011 census

Research question, definitions and content

Research question:
What is the value of census data for population projection?

Definitions:
Census = complete enumeration of the population at their place of “usual” residence every ten or five years
Most questions refer to census date but one or two refer to behaviour over a prior time interval (e.g. migration question, fertility questions, mortality of household members)

Estimation and projection:
Model the development of the population forward in time from a benchmark set by the census
Population attributes are usually quite simple: age, sex and something else

Examples of the contribution of the census to component estimates
Fertility rates by ethnic group
Ethnic mixing matrix
Survival probabilities by ethnic group
Internal migration probabilities by ethnic group
Immigration and emigration estimates by ethnicity

Beyond 2011
What do we need to do if there is not another census?
Here is an all group summary of the different projections we have run for ethnic groups.

Table: Total populations of the UK, 2001-2051: the 2008-based National Population Projections and five ethnic group projections (populations in millions)

<table>
<thead>
<tr>
<th>Year</th>
<th>NPP 2008</th>
<th>BENCH-EF</th>
<th>BENCH-ER</th>
<th>TREND-EF</th>
<th>UPTAP-ER</th>
<th>UPTAP-EF</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>59.1</td>
<td>59.1</td>
<td>59.1</td>
<td>59.1</td>
<td>59.1</td>
<td>59.1</td>
</tr>
<tr>
<td>2010</td>
<td>70.9</td>
<td>63.6</td>
<td>59.5</td>
<td>71.9</td>
<td>67.9</td>
<td>72.3</td>
</tr>
<tr>
<td>2020</td>
<td>77.1</td>
<td>63.1</td>
<td>55.1</td>
<td>77.7</td>
<td>68.3</td>
<td>76.5</td>
</tr>
<tr>
<td>2030</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2040</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2050</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure: Trends in the UK population, ONS 2008-based projections and five ethnic group projections, 2001 to 2051

Projection Specifications
BENCH Benchmark projection using constant 2001-2 component rates, probabilities and flows
TREND Trend projection using estimated 2001-7 or 2001-8 component rates, probabilities and flows; component rates, probabilities and flows thereafter aligned to NPP 2008 assumptions
UPTAP Understanding Population Trends and Processes projection using revised assumptions
EF Emigration flows model
ER Emigration rates model
What do census data have to do with population projection?

Fertility indicators by ethnic group ...

Vital Statistics
- Geography: national to local authority
- Age: 5 year detail
- Ethnicity: no information

Sample
- Geography: national & region
- Age: (SYA & ) 5 year detail
- Ethnicity: (n) 8 & 5 groups (2001)

Census input

Populations
- Geography: national to local
- Age: (SYA & ) 5 year detail
- Ethnicity: up to 16 groups (2001)
Fertility

Estimate fertility by ethnic group ...


- TFRs for ‘intermediate’ groups estimated from TFR for all women using Child : Woman Ratios (CWRs) & the Sample of Anonymised Records (SARs) ...

\[
\text{TFR}(e) = \text{TFR}(aw) * \left( \frac{\text{CWR}(e)}{\text{CWR}(aw)} \right)
\]

Where: \( \text{TFR}(aw) = 1.83 \) in 1991 & 1.63 in 2001

Ethnic mixing matrix

Mothers can give birth to children who have different ethnicities

- Sending percentages
- % of children under one with ethnic group mothers by ethnicity of child

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>WHITE</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Other mixed</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>總計</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Source: Computed by the authors using a 2001 Census Commissioned table.

Notes: The table displays sending percentages, i.e., the percentages of children under one born to mothers of each ethnicity classified by the ethnicity they were assigned in the census. The mother’s ethnicity is represented in the columns and the child’s ethnicity in the rows.
Mortality

Life expectancy at birth 2001

<table>
<thead>
<tr>
<th>Ethnic group</th>
<th>Mean $e_x$ Women</th>
<th>Mean $e_x$ Men</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese</td>
<td>82.1</td>
<td>78.1</td>
</tr>
<tr>
<td>Other White</td>
<td>81.3</td>
<td>76.9</td>
</tr>
<tr>
<td>Other Ethnic</td>
<td>81.5</td>
<td>76.2</td>
</tr>
<tr>
<td>All groups</td>
<td>80.5</td>
<td>76.0</td>
</tr>
<tr>
<td>Black African</td>
<td>80.4</td>
<td>76.1</td>
</tr>
<tr>
<td>White British</td>
<td>80.5</td>
<td>75.9</td>
</tr>
<tr>
<td>White-Irish</td>
<td>80.3</td>
<td>74.9</td>
</tr>
<tr>
<td>White-Asian</td>
<td>80.0</td>
<td>75.1</td>
</tr>
<tr>
<td>Indian</td>
<td>79.3</td>
<td>75.5</td>
</tr>
<tr>
<td>Other-Asian</td>
<td>79.5</td>
<td>75.2</td>
</tr>
<tr>
<td>Other Mixed</td>
<td>79.9</td>
<td>74.6</td>
</tr>
<tr>
<td>White-Black African</td>
<td>79.3</td>
<td>74.2</td>
</tr>
<tr>
<td>Black Caribbean</td>
<td>79.1</td>
<td>74.4</td>
</tr>
<tr>
<td>White-Black Caribbean</td>
<td>78.7</td>
<td>73.4</td>
</tr>
<tr>
<td>Other Black</td>
<td>78.5</td>
<td>73.4</td>
</tr>
<tr>
<td>Bangladeshi</td>
<td>77.7</td>
<td>72.7</td>
</tr>
<tr>
<td>Pakistani</td>
<td>77.2</td>
<td>72.1</td>
</tr>
</tbody>
</table>
### Table: Sub-national migration flows for ethnic groups, Indian ethnic group, 2001 Census

<table>
<thead>
<tr>
<th>Zone</th>
<th>Zone name</th>
<th>Origin</th>
<th>Total migrating within UK</th>
<th>Surviving migrants within UK</th>
<th>Probability of out-migration from area*</th>
<th>Survivors in Rest of UK</th>
<th>Probability of out-migration from RUK*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>City of London + Westminster</td>
<td>405</td>
<td>3</td>
<td>1,015</td>
<td>4,463</td>
<td>0.120226</td>
<td>872</td>
</tr>
<tr>
<td>67</td>
<td>Leeds</td>
<td>1,134</td>
<td>1</td>
<td>11,322</td>
<td>11,859</td>
<td>0.045253</td>
<td>707</td>
</tr>
<tr>
<td>155</td>
<td>Northern Ireland</td>
<td>385</td>
<td>2</td>
<td>180</td>
<td>1,399</td>
<td>0.055575</td>
<td>48</td>
</tr>
</tbody>
</table>

**Notes:** * given survived in the UK. **Figures may not sum precisely to column or row totals because of rounding for presentation purposes.

**Source:** Authors' calculations based on Commissioned Table CO528, 2001 Census, Crown Copyright and census migration statistics and population data from ONS, GROS and NISRA.

---

**Arithmetic to work out surviving stayers**

\[
\text{Surviving stayers} = \text{Total surviving within UK} - \text{Surviving out-migrants} - \text{Total out-migrants within UK}
\]

**Figure:** Migration probabilities for Leeds, by ethnic group, 2000-1

**Source:** Authors’ calculations based on census migration and population data from ONS, GROS and NISRA.
Emigration estimates
Immigration estimates

International migration

But the census information is dated: migration origins change 2001-2007

Age profiles

International migration
Location quotients in 2001 and 2051 Black Caribbean

Location quotients in 2001 and 2051 Pakistani
Location quotients in 2001 and 2051 Chinese

BEYOND 2011

Drivers of change

*Increasing cost of the traditional census*: 2001 = £350m, 2011 = £500m
% growth of 5.2% pa compared with CPI = 2.1% so real price increase of 3.1%
Digital technology is not yielding the benefits expected

*Increasing non-response*: despite the legal requirement to complete, completion rates are falling

*We need more timely information*: waiting two years for results which then remain static is seen as less and less acceptable

*We are living more complex lives* in terms of mobility and time spent at different locations: planning needs to know about lots of different populations (person-time) rather than just one “usually resident” population

*We already collect an enormous amount of digital information about individuals*: using it again would seem sensible (as long as there are proper safeguards and general consent). In the commercial world businesses use customer information in innovative and customer friendly ways (Tesco’s Clubcard, Amazon’s You might also like)
BEYOND 2011

Some countries are already replacing the traditional census:

**France** holds a *rolling census* with 20% enumeration each year so that the workload is spread out over 5 years. Complex statistical interpolation and extrapolation models are then used to estimate the small area population.

The **United States** has shrunk its census to a *short form*. Other characteristics are gathered through a large *American Community Survey*. Cumulating the ACS over 5 years gives a big enough sample to make small area estimates.

The **Netherlands** has a Population Register which can be used to count the population at yearly intervals and then further attributes are added from a set of large surveys.

**Finland** (and other Nordic countries) have joined together their population, administrative, business, property registers to create a *comprehensive database* of the population. The census is just an annual set of tables from this database.

The ESRC-ONS Rolling Census Scoping Study will explore these and other options.

CONCLUSIONS

We have, I hope, convinced you that census data are a necessary though not sufficient input to population projection models.

They need, of course, to be combined with information on the dynamics of the population which comes from registers, administrative records, cross-sectional and longitudinal surveys.

The biggest challenge if we replace the census will be to make reliable estimates of cross-sectional information and of cross-temporal movement/transition between the states of the population we are interested in. We will need to develop methods to fill in the space-time-attribute cube, recognizing the trade-offs involved.