Evaluation of Fertility Data Collected from Population Censuses

United Nations Statistics Division
Outline

1. Fertility data collected in censuses
   1. Children ever born
   2. Recent births
   3. Age-sex structure of population
   4. Micro data on mothers and own-children

2. Quality assessment
   1. Data collection errors, coverage and completeness
   2. Patterns of average parities and parity distributions
   3. Age-specific fertility rates from data on births
   4. Methods for deriving fertility estimates
   5. Comparing estimates from multiple independent sources
Children ever born (summary birth histories)

- Measure of all live births a woman has had in her lifetime
- Asked to all women age 15 and older
- For every woman the following information is collected:
  - a) the total number of female children she has borne in her lifetime.
  - b) the total number of male children she has borne in her lifetime.
  - c) the number of female children who are surviving
  - d) the number of male children who are surviving
Children ever born

- Recommended question sequence to improve completeness of data:
  1. Total number of sons ever born alive during the lifetime of the woman
  2. Total number of sons living (surviving) at the time of the census
  3. Total number of sons born alive who died before the census date
  4. Total number of daughters ever born alive during the lifetime of the woman
  5. Total number of daughters living (surviving) at the time of the census
  6. Total number of daughters born alive who died before the census date

Source: *Principles and Recommendations for Population and Housing Censuses, Rev.2*, United Nations, 2008
Children ever born – When is it used?

- Widely used for over 50 years both for measures of fertility and for child mortality (next session)
- Very important for countries without or with incomplete birth registration
- Also important for countries with complete birth registration
  - Allows for the study of fertility by detailed socio-economic characteristics
Recent births

- Measure of recent fertility
- Asked to all women age 15 – 50 at the time of the census who reported at least one live birth in their lifetime
- **Preferred question:** Date of birth of last child born alive (day, month and year)
- **Alternative question:** Births in the last twelve months to the woman or in the household
  - More error-prone than exact date of birth, although both are subject to under-reporting
  - Date of birth can be converted to births in last 12 months during data processing (will miss only small percentage of cases in which woman had multiple births in a year)
Fertility data – possible errors

Both methods: enumerator’s error

1. Enumerators’ failure to reach individuals
   a) The not-at-home error: information provided by neighbors
   b) Coverage error: omit an area or forgot to record the answer

2. Recording error
   a) Answer is recorded incorrectly by the enumerator
      E.g., Childless women mis-classified into parity not stated
Children ever born – possible errors

1. Errors because the respondent did not understand the question
   a) Mortality error: reported only children living rather than ever-born
   b) Non-resident error: did not report surviving children living elsewhere
   c) Marriage error: women not reporting her children born from previous marriage or children born out of wedlock

2. Errors because of respondents’ lapse of memory or neglect
   a) Memory error: respondent forgot some children
      ➢ Believed to be more common among older women

3. Age misreporting
   a) Teenage mothers may exaggerate their age
   b) Age misreporting if this results in a systematic over- or under-stating of age
Recent births – possible errors

1. Reference period errors
   a) Uncertain of the exact date of birth relative to the reference period
   b) Incorrectly moving birth into or out of the reference period

2. Births missed because mother not located
   a) Women had a birth recently but died or migrated before the census
   b) Household had a birth recently but the household dissolved before the census
   c) Not significant in most cases, however could become an issue when many deaths occurring in a short period (HIV/AIDS) or when there is significant migration
Standard fertility measures

**Average Parity/Children Ever Born** – average number of children had by women in an age group

**Parity Distributions** – distribution of women in each age group by number of children they have had

**Age Specific Fertility Rates (ASFR)** – indicates the age pattern of fertility in a society

\[ nF_x = \frac{nB_x}{nW_x} \]

- \( nB_x \) = Births to women age \( x \) to \( x+n \) during period
- \( nW_x \) = Mid-period population of women age \( x \) to \( x+n \)

**Total Fertility Rate (TFR)** – number of children a woman would have in her lifetime if she lived her whole life under today’s fertility conditions (ASFRs)

\[ TFR = n \sum nF_x \]
### Census fertility data – what can we get?

<table>
<thead>
<tr>
<th></th>
<th>Parity Distribution</th>
<th>Average Parity</th>
<th>ASFR</th>
<th>TFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children Ever Born</td>
<td>Y</td>
<td>Y</td>
<td>Y*</td>
<td>Y*</td>
</tr>
<tr>
<td>Recent Fertility</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

*With one census under constant fertility, otherwise with two censuses*
Evaluating fertility data using standard fertility measures
CEB – quality assessment (Step 1)

- Initial assessment of data quality and missing values
  - Any missing values in children ever born data?
  - Missing value for any relevant variables? (age of mother, sex of child, survival status of the child)
  - Was imputation, hotdecking or any other method used to clean the data?
    - If so, should have a good understanding of the rules followed
CEB – quality assessment

<table>
<thead>
<tr>
<th>Age group</th>
<th>African</th>
<th>Coloured</th>
<th>Indian/Asian</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>12-14</td>
<td>65.2</td>
<td>53.5</td>
<td>61.4</td>
<td>46.2</td>
</tr>
<tr>
<td>15-19</td>
<td>73.5</td>
<td>63.7</td>
<td>68.8</td>
<td>55.9</td>
</tr>
<tr>
<td>20-24</td>
<td>82.5</td>
<td>78.5</td>
<td>79.1</td>
<td>73.9</td>
</tr>
<tr>
<td>25-29</td>
<td>88.2</td>
<td>87.6</td>
<td>88.0</td>
<td>85.4</td>
</tr>
<tr>
<td>30-34</td>
<td>90.9</td>
<td>91.2</td>
<td>92.2</td>
<td>90.2</td>
</tr>
<tr>
<td>35-39</td>
<td>91.9</td>
<td>92.6</td>
<td>93.5</td>
<td>91.3</td>
</tr>
<tr>
<td>40-44</td>
<td>91.4</td>
<td>92.5</td>
<td>93.3</td>
<td>91.5</td>
</tr>
<tr>
<td>45-49</td>
<td>89.9</td>
<td>91.3</td>
<td>91.9</td>
<td>90.4</td>
</tr>
</tbody>
</table>

Source: *Estimation of fertility from the 2001 South Africa census data*, Tom Moultrie & Rob Dorrington, Centre for Actuarial Research, University of Cape Town
CEB – quality assessment (Step 2)

Tabulation of children ever born

- Number of children should not be grouped, except for the last open category (usually no lower than 9+ or 10+ children)
- Children ever born *not stated* should be distinguished from *no children (parity “0”)*
- Are parities reasonable?
  - Quick rule-of-thumb: maximum parity should be one child every 18 months from age of 12
  - E.g. by exact age 20 (end of 15 – 19 age group) maximum children should be 5

# CEB – quality assessment

### Swaziland 1997 Census - Children Ever Born

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>48,289</td>
<td>15,331</td>
<td>5,761</td>
<td>2,575</td>
<td>1,640</td>
<td>1,075</td>
<td>763</td>
</tr>
<tr>
<td>1</td>
<td>6,687</td>
<td>14,368</td>
<td>6,558</td>
<td>3,326</td>
<td>1,992</td>
<td>1,248</td>
<td>878</td>
</tr>
<tr>
<td>2</td>
<td>1,081</td>
<td>9,100</td>
<td>8,277</td>
<td>4,256</td>
<td>2,612</td>
<td>1,587</td>
<td>1,116</td>
</tr>
<tr>
<td>3</td>
<td>150</td>
<td>3,579</td>
<td>7,059</td>
<td>4,602</td>
<td>3,106</td>
<td>1,811</td>
<td>1,274</td>
</tr>
<tr>
<td>4</td>
<td>35</td>
<td>1,196</td>
<td>4,632</td>
<td>4,535</td>
<td>3,320</td>
<td>2,087</td>
<td>1,474</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>2,382</td>
<td>3,736</td>
<td>3,116</td>
<td>1,980</td>
<td>1,497</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>4115</td>
<td>1,067</td>
<td>2,801</td>
<td>2,915</td>
<td>1,989</td>
<td>1,584</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>54</td>
<td>436</td>
<td>1,694</td>
<td>2,494</td>
<td>1,925</td>
<td>1,603</td>
</tr>
<tr>
<td>8 - 9</td>
<td>0</td>
<td>47</td>
<td>277</td>
<td>1,387</td>
<td>3,074</td>
<td>3,015</td>
<td>2,774</td>
</tr>
<tr>
<td>10+</td>
<td>0</td>
<td>10</td>
<td>68</td>
<td>430</td>
<td>1,285</td>
<td>2,206</td>
<td>2,602</td>
</tr>
<tr>
<td>Unknown</td>
<td>1,331</td>
<td>2,150</td>
<td>1,379</td>
<td>826</td>
<td>603</td>
<td>417</td>
<td>345</td>
</tr>
</tbody>
</table>

Data source: United Nations Demographic Yearbook

United Nations Workshop on Census Data Evaluation for English Speaking African Countries
Kampala, Uganda
12 – 16 November 2012

Parity 6 and 7 are obviously wrong.
Parity 8 and 9 should not have been grouped.
Unknown separate from parity “0”
### CEB – quality assessment

#### Swaziland 1997 Census - Children Ever Born

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>48,289</td>
<td>15,331</td>
<td>5,761</td>
<td>2,575</td>
<td>1,640</td>
<td>1,075</td>
<td>763</td>
</tr>
<tr>
<td>1</td>
<td>6,687</td>
<td>14,368</td>
<td>6,558</td>
<td>3,326</td>
<td>1,992</td>
<td>1,248</td>
<td>878</td>
</tr>
<tr>
<td>2</td>
<td>1,081</td>
<td>9,100</td>
<td>8,277</td>
<td>4,256</td>
<td>2,612</td>
<td>1,587</td>
<td>1,116</td>
</tr>
<tr>
<td>3</td>
<td>150</td>
<td>3,579</td>
<td>7,059</td>
<td>4,602</td>
<td>3,106</td>
<td>1,811</td>
<td>1,274</td>
</tr>
<tr>
<td>4</td>
<td>35</td>
<td>1,196</td>
<td>4,632</td>
<td>4,535</td>
<td>3,320</td>
<td>2,087</td>
<td>1,474</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>337</td>
<td>2,382</td>
<td>3,736</td>
<td>3,116</td>
<td>1,980</td>
<td>1,497</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>115</td>
<td>1,067</td>
<td>2,801</td>
<td>2,915</td>
<td>1,989</td>
<td>1,584</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>54</td>
<td>436</td>
<td>1,694</td>
<td>2,494</td>
<td>1,925</td>
<td>1,603</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>47</td>
<td>277</td>
<td>1,387</td>
<td>3,074</td>
<td>3,015</td>
<td>2,774</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>68</td>
<td>430</td>
<td>1,285</td>
<td>2,206</td>
<td>2,602</td>
</tr>
<tr>
<td>10+</td>
<td>0</td>
<td>68</td>
<td>430</td>
<td>1,285</td>
<td>2,206</td>
<td>2,602</td>
<td>1,584</td>
</tr>
</tbody>
</table>

Total children: 57,581 46,287 37,896 30,168 26,157 19,340 15,910

Total women: 9,454 51,242 87,216 107,218 119,321 101,200 90,637

Proportion with unknown parity should stay constant
Proportion childless should decrease with age
Average parity should increase with age

- Proportion childless: 0.0232 0.04667 0.03639 0.02738 0.02305 0.02156 0.02168
- Proportion childless: 0.83863 0.33122 0.15202 0.08536 0.0627 0.05558 0.04796
- Average parity: 0.16419 1.10704 2.30144 3.55401 4.56172 5.23265 5.69686
CEB – quality assessment

Average parity at age \( x \):

\[
P_x = \frac{B_x}{W_x} = \frac{\sum_{j} j W_{j,x}}{\sum_{j} W_{j,x}}
\]

where

\( B_x \) - number of births

\( W_{j,x} \) - number of women at parity \( j \)
CEB – quality assessment

Average parity, Swaziland 1997, all age groups

Average parity, Gabon 1993

Misreporting? Rising fertility?

Underreporting at higher ages?

Misreporting?
Rising fertility?

Underreporting at higher ages?
The El-Badry Correction

- El-Badry correction is applied to adjust reported data on children ever born any further analysis
- A common problem with CEB data is that enumerators may incorrectly code women of zero parity as “parity unknown” or “parity not stated”
- The El-Badry method corrects for this by apportioning those women with parity ‘reportedly’ unknown between those whose parity is ‘truly’ unknown and those who have no children
- Method is based on assumption that proportion of women whose parity is ‘truly’ unknown does not depend on age

Application

- Check if proportion of women with parity unknown is high and going down with age
- If parity unknown is less than 2% of each age group it is safe to assume that the data are not affected and no correction is needed
Identifying when to use El-Badry method

**Proportion of women with unknown parity * 100%**
Swaziland, Kenya and Burundi

High proportion of women in younger age groups with parity unknown suggests that some women with no children were misclassified and El-Badry correction should be applied.
El-Badry: Step 1

Calculate proportion of women in each age group with
a) parity missing and b) parity = 0

a) Parity unknown:

\[ U_i = \frac{N_{i,u}}{N_i} \]

Where:

- \( U_i \) = proportion unknown in age group
- \( N_{i,u} \) = number unknown in age group
- \( N_i \) = total women in age group

b) Parity 0:

\[ Z_i = \frac{N_{i,0}}{N_i} \]

Where:

- \( Z_i \) = proportion parity 0 in age group
- \( N_{i,0} \) = number parity 0 in age group
- \( N_i \) = total women in age group
### Parity data, Kenya 1989 Census

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>597,560</td>
<td>198,600</td>
<td>59,400</td>
<td>23,120</td>
<td>14,580</td>
<td>11,040</td>
<td>9,560</td>
</tr>
<tr>
<td>1</td>
<td>134,700</td>
<td>224,660</td>
<td>83,140</td>
<td>26,140</td>
<td>13,620</td>
<td>9,460</td>
<td>7,740</td>
</tr>
<tr>
<td>2</td>
<td>38,120</td>
<td>202,300</td>
<td>120,940</td>
<td>38,340</td>
<td>19,180</td>
<td>13,240</td>
<td>9,280</td>
</tr>
<tr>
<td>3</td>
<td>11,120</td>
<td>126,500</td>
<td>150,500</td>
<td>53,880</td>
<td>28,020</td>
<td>17,000</td>
<td>12,440</td>
</tr>
<tr>
<td>4</td>
<td>6,820</td>
<td>59,700</td>
<td>146,500</td>
<td>73,280</td>
<td>37,340</td>
<td>21,400</td>
<td>14,800</td>
</tr>
<tr>
<td>5</td>
<td>1,740</td>
<td>33,720</td>
<td>102,300</td>
<td>87,720</td>
<td>48,140</td>
<td>28,980</td>
<td>18,560</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>12,480</td>
<td>58,980</td>
<td>83,580</td>
<td>56,520</td>
<td>35,260</td>
<td>26,280</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>57,180</td>
<td>91,800</td>
<td>56,240</td>
<td>41,260</td>
<td>28,640</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>64,740</td>
<td>56,560</td>
<td>42,700</td>
<td>32,920</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>40,780</td>
<td>39,480</td>
<td>33,000</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>26,840</td>
<td>32,240</td>
<td>27,920</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>14,920</td>
<td>22,840</td>
<td>21,920</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8,280</td>
<td>14,660</td>
<td>14,720</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3,740</td>
<td>7,900</td>
<td>8,920</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2,180</td>
<td>4,080</td>
<td>4,900</td>
</tr>
<tr>
<td>15+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3,160</td>
<td>5,400</td>
<td>7,180</td>
</tr>
<tr>
<td></td>
<td><strong>U</strong></td>
<td><strong>402,780</strong></td>
<td><strong>147,540</strong></td>
<td><strong>61,920</strong></td>
<td><strong>31,580</strong></td>
<td><strong>21,480</strong></td>
<td><strong>16,060</strong></td>
</tr>
<tr>
<td></td>
<td>Total women</td>
<td>1,192,840</td>
<td>1,005,500</td>
<td>840,860</td>
<td>574,180</td>
<td>451,580</td>
<td>363,000</td>
</tr>
<tr>
<td></td>
<td><strong>Ui</strong></td>
<td>0.338</td>
<td>0.147</td>
<td>0.074</td>
<td>0.055</td>
<td>0.048</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td><strong>Zi</strong></td>
<td>0.501</td>
<td>0.198</td>
<td>0.071</td>
<td>0.040</td>
<td>0.032</td>
<td>0.030</td>
</tr>
</tbody>
</table>

United Nations Workshop on Census Data Evaluation for English Speaking African Countries
Kampala, Uganda
12 – 16 November 2012
El-Badry: Step 2

El Badry correction fit, Kenya 1989 census

United Nations Workshop on Census Data Evaluation for English Speaking African Countries
Kampala, Uganda
12 – 16 November 2012
El-Badry: Step 3

- Regress $U_i$ on $Z_i$ (in excel can use SLOPE and INTERCEPT functions)

In our example, get intercept ($\beta$) of .0275, suggesting 2.7% of data of each age group is truly missing.

To correct data:

Parity truly missing = $U'_i = N_i \times \beta$

Parity 0 =

$N'_{i,0} = N_i (Z_i + U_i - \beta)$
### El-Badry: Step 4

#### Revised figures for women with unknown and 0 parity, Kenya 1989 census with El Badry correction

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ni,0</td>
<td>597,560</td>
<td>198,600</td>
<td>59,400</td>
<td>23,120</td>
<td>14,580</td>
<td>11,040</td>
<td>9,560</td>
</tr>
<tr>
<td>Ui</td>
<td>402,780</td>
<td>147,540</td>
<td>61,920</td>
<td>31,580</td>
<td>21,480</td>
<td>16,060</td>
<td>13,540</td>
</tr>
<tr>
<td>Total women</td>
<td>1,192,840</td>
<td>1,005,500</td>
<td>840,860</td>
<td>574,180</td>
<td>451,580</td>
<td>363,000</td>
<td>292,320</td>
</tr>
<tr>
<td>Ui</td>
<td>0.338</td>
<td>0.147</td>
<td>0.074</td>
<td>0.055</td>
<td>0.048</td>
<td>0.044</td>
<td>0.046</td>
</tr>
<tr>
<td>Zi</td>
<td>0.501</td>
<td>0.198</td>
<td>0.071</td>
<td>0.040</td>
<td>0.032</td>
<td>0.030</td>
<td>0.033</td>
</tr>
</tbody>
</table>

\[
U'_i = N_i \times \beta
\]

\[
N'_{i,0} = N_i (Z_i + U_i - \beta)
\]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>32,803</td>
<td>27,651</td>
<td>23,124</td>
<td>15,790</td>
<td>12,418</td>
<td>9,983</td>
<td>8,039</td>
</tr>
<tr>
<td></td>
<td>967,537</td>
<td>318,489</td>
<td>98,196</td>
<td>38,910</td>
<td>23,642</td>
<td>17,118</td>
<td>15,061</td>
</tr>
</tbody>
</table>
Recalculation of average parity after El-Badry

- If the el-Badry method has been applied, average parities should be calculated excluding the remaining ("true") number of women with unknown parity from the denominator.
  - This will increase the average parities by $1/(1+\beta)$ because women formerly considered missing are now classified as parity 0.

- When missing data is more than 2% but the correction is not applied (e.g. due to violation of linearity), women of unknown parity should be included in the denominator.
  - This will lead to underestimation of average parity because the unknown parities are functionally treated as parity 0.
El-Badry: revised parities

Average parity, reported and corrected by el-Badry method, Kenya, 1989 census

- Reported
- Corrected by el-Badry method
CEB checks – Parity distribution of women age 45 - 49

- High level of parity 0 in 1950 and 1970 censuses: possibly groups “not stated” and “0” parity combined. No separate groups unlike as in the 1980 census.

- Flat curve: probably some form of misreporting, seems to be improving over time

- Mexican fertility survey: shape of the curve more plausible (small sample size)

CEB –Checks – Parity distribution of women age 45 - 49

Data source: United Nations Demographic Yearbook
• Simple test for quality of reporting among older women

• Assumes all childbearing at age 25

• Year in time = census year – (age – 25)

• Thailand example: 1960 and 1970 censuses - an increase in fertility

• Erroneous data from 1980 census (conclusion was reached after comparing with data from other surveys)

CEB - Additional Checks - Cohort analysis of mean number of children ever born

**Time plot of mean children ever born, Botswana 1971 - 2001**

Data source: United Nations Demographic Yearbook

United Nations Workshop on Census Data Evaluation for English Speaking African Countries
Kampala, Uganda
12 – 16 November 2012
CEB – Additional checks – multiple sources of data


United Nations Workshop on Census Data Evaluation for English Speaking African Countries
Kampala, Uganda
12 – 16 November 2012
## CEB – Additional checks – multiple sources of data

Malawi census form for 2008 – fertility section

### Table: CEB – Additional checks – multiple sources of data

Malawi census form for 2008 – fertility section

<table>
<thead>
<tr>
<th>P30. How many children were born alive to (NAME)?</th>
<th>P31. Among those children, how many are still alive?</th>
<th>P32. How many live births during the last 12 months?</th>
<th>P33. Among those children born in the last 12 months, how many are still alive?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

*Note: Write number of boys and girls here.*

**MEMBERS. IF THE PERSON IS THE LAST MEMBER OF THE HOUSEHOLD, PROCEED TO SECTION D**
Recent births – quality assessment

☐ Initial assessment

- Any missing values in data? (month/date/year of birth)
  - Missing data for any relevant variables? (age of mother, sex of child, survival status of the child)
- Is distribution of reported birth dates reasonable?
- If possible, compare with civil registration data on live births
Recent births – quality assessment – missing and inconsistent data

Table 2.9 Distribution of women aged 12 to 49 by imputation flag for response to question on year of last child’s birth

<table>
<thead>
<tr>
<th></th>
<th>No imputation missing response</th>
<th>Logical imputation from non-missing response</th>
<th>Hotdeck applied to missing response</th>
<th>Hotdeck applied to non-missing response</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women (per cent)</td>
<td>6560661</td>
<td>604260</td>
<td>391548</td>
<td>734257</td>
<td>8455728</td>
</tr>
<tr>
<td></td>
<td>77.6</td>
<td>7.1</td>
<td>4.6</td>
<td>8.7</td>
<td>77.6</td>
</tr>
</tbody>
</table>

Source: *Estimation of fertility from the 2001 South Africa census data*, Tom Moultrie & Rob Dorrington, Centre for Actuarial Research, University of Cape Town
Recent births – quality assessment - sex ratio

Sex ratio at birth, births in 12 months preceding census, selected censuses

Data source: United Nations Demographic Yearbook

United Nations Workshop on Census Data Evaluation for English Speaking African Countries
Kampala, Uganda
12 – 16 November 2012
Recent births – quality assessment – age specific fertility rates (ASFR)

Age Specific Fertility Rate (ASFR)

\[
\frac{nBx}{nWx} = \text{Births to women age } x \text{ to } x+n \text{ during period} \\
\frac{nFx}{nWx} = \text{Mid-period population of women age } x \text{ to } x+n
\]

### Malawi, census June 2008

<table>
<thead>
<tr>
<th>Age group (years)</th>
<th>Births in 12 months preceding census</th>
<th>Total women in age group</th>
<th>ASFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.5 - 19.5</td>
<td>70,737</td>
<td>699,155</td>
<td>0.10117</td>
</tr>
<tr>
<td>19.5 - 24.5</td>
<td>169,406</td>
<td>596,363</td>
<td>0.28407</td>
</tr>
<tr>
<td>24.5 - 29.5</td>
<td>130,331</td>
<td>539,482</td>
<td>0.24159</td>
</tr>
<tr>
<td>29.5 - 34.5</td>
<td>79,232</td>
<td>517,345</td>
<td>0.15315</td>
</tr>
<tr>
<td>34.5 - 39.5</td>
<td>43,747</td>
<td>374,526</td>
<td>0.11681</td>
</tr>
<tr>
<td>39.5 - 44.5</td>
<td>15,956</td>
<td>276,264</td>
<td>0.05776</td>
</tr>
<tr>
<td>44.5 - 49.5</td>
<td>5,599</td>
<td>224,100</td>
<td>0.02498</td>
</tr>
</tbody>
</table>

Are births be classified by age of mother at birth of her child or by age of mother at the survey/census date? If not known, assume the latter, almost universally, in censuses, data are classified by age of mother at time of census. In this case, ASFRs are shifted by ½ year as mothers were ½ year younger at the time of birth.
Recent births – quality assessment – comparing ASFRs

Data source: United Nations Demographic Yearbook and DHS STATcompiler
Recent births – quality assessment – comparing TFRs

Total fertility rate

\[ \text{TFR} = 5 \sum_{x} F_x \]

Malawi TFR comparison

<table>
<thead>
<tr>
<th>Age group</th>
<th>2004 DHS</th>
<th>2008 Census</th>
<th>2010 DHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 - 19</td>
<td>0.810</td>
<td>0.506</td>
<td>0.760</td>
</tr>
<tr>
<td>20 - 24</td>
<td>1.465</td>
<td>1.420</td>
<td>1.345</td>
</tr>
<tr>
<td>25 - 29</td>
<td>1.270</td>
<td>1.208</td>
<td>1.190</td>
</tr>
<tr>
<td>30 - 34</td>
<td>1.110</td>
<td>0.766</td>
<td>1.030</td>
</tr>
<tr>
<td>35 - 39</td>
<td>0.815</td>
<td>0.584</td>
<td>0.810</td>
</tr>
<tr>
<td>40 - 44</td>
<td>0.400</td>
<td>0.289</td>
<td>0.410</td>
</tr>
<tr>
<td>45 - 49</td>
<td>0.175</td>
<td>0.125</td>
<td>0.165</td>
</tr>
<tr>
<td>TFR</td>
<td>6.05</td>
<td>4.90</td>
<td>5.71</td>
</tr>
</tbody>
</table>

United Nations Workshop on Census Data Evaluation for English Speaking African Countries
Kampala, Uganda
12 – 16 November 2012
Estimating fertility from data collected in censuses

- To obtain new estimates of fertility
- To compare estimates from the current census with estimates available from other sources e.g. surveys
Lesotho, fertility estimates from different sources

Lesotho, TFR

Year


TFR


Lesotho, fertility estimates from different sources:

- WPP 2010
- WFS1977 (CBS, Lesotho)
- Censuses 1986/1996 (Rel. Gompertz, in 2001 LesoDemoSurv)
- WFS1977 (CBS, Lesotho)
- Makatjane, TJ / 1991 Lesotho Demographic and Health Survey (Roma)
- 2002 LesoRepHealthSurv (Rel. Gompertz and Reported)
- DHS2009d
- 2006 Census
- R10
- Censuses 1986/1996 (Reported, in 2001 LesoDemoSurv)
- Censuses 1986/1996 (Brass PF ratio, in 2001 Leso DemoSurv)
- 2001 LesoDemoSurv (Reported, in 2004 LDHS)
- 2004 LDHS
- Census 2006

United Nations Workshop on Census Data Evaluation for English Speaking African Countries
Kampala, Uganda
12 – 16 November 2012
Methods for estimating fertility

- Interpolation of average parities (Mortara, 1949)
- Brass P/F method and its variations and extensions, e.g. Arriaga (1983), Relational Gompertz model
- Methods based on population structure: Reverse Survival and Own Children Method
- Methods based on data from two or several censuses: Arriaga (1983), synthetic relational Gompertz model, parity increments
Interpolation and backdating average parities

Average parity at ages \( x, x+n \) by definition:

\[
n P_x = \int_{x}^{x+n} F(a) da
\]

where \( F \) is cohort cumulative fertility function.

- By using interpolation one can compute age-specific fertility rates from average parities, \( P \), assuming that fertility was more or less constant before the census.
- For ages with completed fertility, e.g. age > 45, we can assume that \( P \approx TFR \), total fertility for a given cohort.
- By plotting \( P \approx TFR \) at years defined by the census date and mean age at childbearing, one can produce estimates of historical TFR trends (Feeney, 1991, see slide presented before).
- Software: FERTCB procedure, Mortpak, UN
The P/F ratio method: Rationale

- The P/F method aims to balance out the strengths and weaknesses of CEB and recent fertility data by comparing:
  1. Cumulative fertility equivalent derived from recent fertility data “F” (trusting the age pattern of fertility but not level)
  2. Life-time average parities “P” (trusting the overall level but not the age distribution)
- The method is typically used to adjust estimates of current fertility level (computed from data on recent births or from incomplete civil registration)
- The method is also used to assess the quality of CEB data and, sometimes, the age reporting of the mother
- Works well if fertility was constant before the census (improbable now); no severe problems with the data

P/F Method: Data requirements

- Total number of children ever born by 5-year age group of mother
- Recent fertility by 5-year age group of mother, measured either by:
  - Births in past year question on census
  - Births registered in year of census from vital registration
- Total number of women in each 5-year age group
P/F Method: Assumptions

- Assumptions:
  - Mis-reporting of current fertility is constant across all age groups
  - Increasing under-reporting of parity (children ever born) by age of women
  - Constant fertility (most important for youngest age groups – up to 35 or so)
    - Can be relaxed through a modification of the original P/F ratio method that uses two consecutive censuses or fertility rates derived from vital registration or another data source
P/F Method: Computational procedure

Procedure described here follows Arriaga (1983) which is implemented in MortPak

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age Group</td>
<td>p(i)</td>
<td>f(i)</td>
<td>p*(i)</td>
<td>f*(i)</td>
<td>P(i)</td>
<td>F(i)</td>
</tr>
</tbody>
</table>

Average CEB as shown
ASFRs as shown
CEB transformed into age-specific rates
ASFR adjusted for time of census enumeration
Cumulated P(i) and F(i)
Adjustment factor for fertility rates, usually ages groups 20-24, 25-29 or 30-35 as the most reliable
P/F method: Interpretation

- Typical “look” of P/F ratios:
  - With perfect data, ratio should be the same for all age groups and close to 1
  - In practice, ok if ratios for 20 – 24, 25 – 29 and (less important) 30 – 34 are close
  - Typically, P/F ratio will decrease with women’s age
  - Deviation from the above typical pattern: indicates either violations of the assumptions or different patterns of under-reporting
Example in MortPak: Malawi 2008 Census

```
<table>
<thead>
<tr>
<th>Age Group of Woman</th>
<th>Children Ever Born</th>
<th>Fertility Pattern (A.5.F.P.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 - 20</td>
<td>0.283</td>
<td>0.1110</td>
</tr>
<tr>
<td>20 - 25</td>
<td>1.532</td>
<td>0.2450</td>
</tr>
<tr>
<td>25 - 30</td>
<td>2.049</td>
<td>0.1900</td>
</tr>
<tr>
<td>30 - 35</td>
<td>4.185</td>
<td>0.1500</td>
</tr>
<tr>
<td>35 - 40</td>
<td>5.214</td>
<td>0.1470</td>
</tr>
<tr>
<td>40 - 45</td>
<td>6.034</td>
<td>0.0720</td>
</tr>
<tr>
<td>45 - 50</td>
<td>8.453</td>
<td>0.0320</td>
</tr>
</tbody>
</table>
```
In the present case the adjustment factors for age groups 20-25, 25-30 and 30-35 are fairly consistent leading to similar levels of adjusted TFRs.
P/F Method: Interpretation

- Example 1: a **declining trend in the P/F ratios** by age of women could indicate that 1) fertility has been increasing or 2) that reported data on children ever born suffer from progressively increasing omissions of children as age of women increases.

- Example 2: **large fluctuations in the P/F ratios** may reflect either differential coverage by age or selective age misreporting by women.

- Example 3: a **rising trend in the P/F ratios** by age of women indicates that fertility could have been decreasing in the past.
Variants on the P/F method

- P/F method for first births – not affected by fertility decline through higher-parity control
- Two-census methods, deriving age schedule of fertility from the two censuses or an additional source (such as vital registration)
  - Can be implemented in MortPak FERTPF by adding optional data for second census
- The Relational Gompertz model uses the same data as the P/F model, but
  - Does not require an assumption of constant fertility
  - Compares/replaces recent fertility data with model fertility schedules to check accuracy
  - Relies on parity data for all age groups (not just younger ones)

Sources: *Estimation of fertility from the 2001 South Africa census data, Manual X, and IUSSP Demographic Estimation*
Relational Gompertz model

- An improved and more versatile version of the Brass P/F method with the same input data
- Shape of fertility distribution adheres to Gompertz relational model
- Level is estimated from average parities
- Robust
- Can be used for smoothing and extrapolation of fertility schedule
- Can be used with different standard patterns

Reverse Survival

- Census population by age and sex is 15-year back projected (reverse survived)
- TFR for years y-1, y-2, ... y-15 computed to match births obtained by reverse survival

Assumptions

- Population by age and sex is free of errors
- Estimates of mortality are available for the period before census
- Reasonably good assumptions can be made about age pattern of fertility (PASFR)

Fertility estimates by Reverse Survival for Myanmar

Myanmar, Total Fertility Rate

Year

TFR

United Nations Workshop on Census Data Evaluation for English Speaking African Countries
Kampala, Uganda
12 – 16 November 2012
Own-children method

- Based on the same idea as reverse survival
- Produces estimates of both TFR and age pattern of fertility

Data requirements

- Distribution of own children by age and by age of mother
- Estimates of mortality for the period before census

Software


Reference

**Step 1**

Obtain distribution of own children by age and by age of mother:

<table>
<thead>
<tr>
<th>Age of mother</th>
<th>Number of children, by age of child</th>
<th>Number women</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>13  7  0  2  2  2  1  4  3  4  2  1  3  1  3  3  755</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>12  3  0  2  0  2  1  1  1  0  0  2  0  0  1  696</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>23  16  6  1  0  0  0  0  0  1  0  0  0  0  0  686</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>58  36  17  3  0  3  1  2  0  0  0  0  0  0  0  706</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>66  46  24  13  11  1  3  1  0  0  0  1  0  0  1  538</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>77  55  45  33  19  12  2  1  0  2  2  1  1  1  0  602</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>78  71  56  47  48  17  7  5  3  0  1  2  1  1  0  488</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>84  80  76  73  46  26  18  15  3  0  0  0  0  1  534</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>84  80  84  61  53  29  24  7  9  1  2  0  2  1  1  0  468</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>93  63  78  72  56  48  45  34  17  9  8  3  0  1  1  1  411</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>91  84  87  83  69  71  55  52  31  21  5  5  2  1  0  464</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>73  67  65  70  66  70  61  55  41  24  17  11  1  1  2  0  393</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>58  61  70  58  63  79  64  64  47  28  27  16  11  5  2  1  339</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>83  71  77  81  94  80  87  91  80  60  42  34  16  8  3  2  442</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>48  58  52  59  68  64  77  75  61  66  48  50  23  23  6  4  330</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>46  60  70  62  82  86  86  82  74  69  50  45  34  12  8  403</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>42  39  42  36  44  44  55  66  63  56  57  46  43  24  12  8  243</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>45  50  67  54  66  65  73  82  79  91  78  64  63  66  38  30  343</td>
<td></td>
</tr>
</tbody>
</table>

Usually requires tabulations of microdata. Algorithms for matching mothers and own children can be fairly complicated.

**Step 2**

Apply reverse survival techniques to the distribution obtained at the previous step to estimate shape and level (TFR) of fertility in the last 15 year.
Own-children method: FERT software

United Nations Workshop on Census Data Evaluation for English Speaking African Countries
Kampala, Uganda
12 – 16 November 2012
Fertility Estimates by Own-Children Method, Kenya

Using DHS microdata with recorded information on mothers. Not using matching algorithm for linking mothers and own children.
References


*Estimation of fertility from the 2001 South Africa census data*, Tom Moultrie & Rob Dorrington, Centre for Actuarial Research, University of Cape Town

IUSSP Tools for Demographic Estimation (in progress, see chapter on fertility)
http://demographicestimation.iussp.org/

*Manual X: Indirect Techniques for Demographic Estimation*, 1983, United Nations (see Chapter 2)


MortPak manual (accompanies software)

East-West Center (www.eastwestcenter.org) (software)
THANK YOU....