Census Data Evaluation
The Case of the 2008 National Population and Housing Census of Liberia
Outline

- Introduction/Background
- Objectives of Census Data Analysis
- Data Evaluation
- Smoothing Methods Used
- Conclusion
Introduction/Background

- Quality of census data is an important feature as it increases its reliability;
- Despite this important aspect, census data all over the world are prone to errors;
- The common errors are age and sex misreporting and digit preference;
- These errors are common in developing countries than in developed countries;
- Hence, it is imperative to evaluate age-sex data from census before they are used.
This presentation provides a brief evaluation of the age-sex distribution of the population of Liberia using the 2008 NPHC data to assess the accuracy of age reporting.
Objectives of the Census Evaluation Program

- To detect irregularities
- To assess the accuracy of age reporting
- To detect measures of digit preference
- To correct the irregularities in age and sex data as well as for digit preference.
Data Evaluation

- To measure the accuracy of the 2008 census data the following indices were calculated:
  - Age ratio scores for males
  - Age ratio scores for females
  - Sex ratio scores
- A combination of these scores give what is known as an Accuracy index of 31.8.
## Data Evaluation cont’d)

<table>
<thead>
<tr>
<th>Index</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age ratio score for males</td>
<td>4.1</td>
</tr>
<tr>
<td>Age ratio score for females</td>
<td>8.8</td>
</tr>
<tr>
<td>Sex ratio score</td>
<td>6.3</td>
</tr>
<tr>
<td>Age-sex accuracy index</td>
<td>31.8</td>
</tr>
</tbody>
</table>
Data Evaluation (cont’d)

- The resulting age-sex accuracy index has a value of 31.8.
- In demographic analysis, age reporting is deemed satisfactory if the data has an accuracy index less than 20.
- The value of the age-sex accuracy index of 31.8 indicates that age reporting was inaccurate.
- Because of this, attempts were made to correct the irregularities in the age-sex data as well as for digit preference.
Data Evaluation (cont’d)

Age misreporting is another errors that census data suffer from

- This may be due to the following causes:
  - Ignorance of correct age;
  - Carelessness in reporting and recording;
  - A general tendency to state age figures ending in certain preferred digits (digit preference);
Data Evaluation (cont’d)

- A tendency to exaggerate length of life at advances ages;
- Possibly subconscious aversion to certain numbers; and
- Mis-statements arising from motives of an economic, social, political, or pure personal character.
Data Evaluation (cont’d)

- In order to get a broad picture about the magnitude of age preference, Whipples, Myers and the Bachi indices were computed separately for males and females and for both sexes.
- The respective values are 1.42, 19.6, and 11.8 for both sexes.
- All three indices support the earlier argument which indicates that age reporting was inaccurate.
Data Evaluation (cont’d)

- **Summary indices of age mis-reporting (Liberia)**

<table>
<thead>
<tr>
<th>Index</th>
<th>Male</th>
<th>Female</th>
<th>Both Sexes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whipples</td>
<td>1.36</td>
<td>1.48</td>
<td>1.42</td>
</tr>
<tr>
<td>Myers</td>
<td>17.7</td>
<td>21.4</td>
<td>19.6</td>
</tr>
<tr>
<td>Bachi</td>
<td>10.5</td>
<td>13.51</td>
<td>1.8</td>
</tr>
</tbody>
</table>
Data Evaluation (cont’d)

- Results of digit preference give the impression of higher concentration of age reporting around digit 0, 2, 5, and 8; and in ages 10, 20, 30, 40, 50 and 60, confirming what indicating that both males and females aged below 60 years misreported their ages, thus calling for age smoothing.
Data Evaluation (cont’d)

- All the calculated indices point to the fact that the census data show some level of unreliability/inaccuracy and needed to be corrected/adjusted.
Methods Used

- The PAS module known as AGESMTH spreadsheet [4] was used to correct the irregularities in age and sex data as well as for digit preference.
- This module employs five smoothing techniques including the Carrier Farrag;
- Karup-King Newton;
- United Nations;
- Strong Moving average; and
- the Arriaga light smoothing which adjusts the population distribution in the first 10-year age groups and incorporates the outcome into 5-year age groups;
- The Arriaga method was used to generate smoothed population for Liberia.
Conclusion

- There were irregularities in the age-sex data resulting from digit preference and age mis-statement.
- These irregularities were detected through the calculation of relevant indices.
- Therefore, it was not possible to use such data until they were corrected.
- The PAS module known as AGESMTH spreadsheet [4] was used to correct the irregularities in age and sex data as well as for digit preference.
Conclusion

THANK YOU!!!