Did you sleep here last night?
The impact of the household definition in sample surveys: a Tanzanian case study

Tiziana Leone, Ernestina Coast (LSE)
Sara Randall (UCL)

Funded by ESRC survey methods initiative
Outline

- Rationale
- Data and methods
- Fieldwork experience
  - Scenarios
- Thinking creatively with the DHS
- Possibly how to fix it
- Discussion and a few thoughts
Do household definitions matter?

- More variables being added in ‘household section’
  - Way of measuring wealth / poverty / access to facilities which influence health
  - New level of analysis / explanation
- More use (researchers & policy makers) made of publicly available data
- Recognition of importance of society’s basic unit as influence upon members’ well-being
- Increasing use of ‘indicators’ based on household data (e.g. MDGs, asset indicators)
- Increasing importance of poverty mapping which uses household level data
How could the definition influence the outcome?

- Household size
- Gender imbalance
- Sex of head of household
- All other characteristics of head of hh
  - Education
  - Occupation etc
- Age structure
- Undercount of “special” populations
- Measures of poverty
  - assets
Household definition in developing countries’ surveys

• Much more standardised (still some local variations)
  • WFS left more space for interpretation
• Little variation between core questionnaires and those used by countries
• Little development over time
• Emphasis on comparability across time and space
Aims and objectives

• Understand the impact of the household definition on:
  – Key demographic indicators
  – Policy making/interventions

• Investigate flexible data collection and analysis
Definitions

- **DHS**: “for the purpose of the 2004-5 TDHS a household is defined as a person or group of persons, related or unrelated who live together and share a common source of food”

- **2002 census** “For the purpose of the 2002 population and housing census a "private household" was a group of persons who lived together and shared living expenses. Usually these were husband, wife, and children. Other relatives, boarders, visitors and servants were included as members of the household, if they were present in the household on census night. If one person lived and ate by himself/herself, then he/she was a one-person household even if he/she stayed in the same house with other people (these cases were more prevalent in the urban areas). Household members staying in more than one house were enumerated as one household if they ate together."
Data and Methods

1. Primary in-depth (n=52) case study interviews with Tanzanians in four different settings.
   - Mix of cognitive interviewing and in-depth Household grid sheet-flexible data collection-573 individuals
     1. Longido in prevalently Maasai area (9 ‘households’)
     2. Urban Dar es Salaam (23 ‘households’)
     3. South Tanzania Rufiji (20 ‘households’)

2. 2004 Tanzanian Demographic and Health Survey (DHS) (n=9735 households)- Household and individual level recodes
   - Scenario analysis-key demographic indicators
1. Summary of fieldwork experience

- Complex cultural traditions around eating meals and sleeping arrangements
- **Maasai** have interdependent groups that are split up in surveys but considered by themselves to be one economic unit of production and consumption
- **Dar es Salaam urban**: very high mobility between households of children and young people
- **Rufiji** Straightforward livelihoods with extremely complex ways of living: subsistence economy with several members contributing to household finances
  - No local word for a household – which suggests not an easy concept
Modelling definition differences

- ‘Translated’ the household grid interviews into SPSS dataset
- We allowed for extra columns to include variables such as:
  - Would this person make it into DHS
  - Would this person make it into Census
- Created simple demographic indicators such as
  - Dependency ratio
  - % female headed household
  - Household size
  - Head of Household education level
## Fieldwork scenarios:

<table>
<thead>
<tr>
<th></th>
<th>Number of households</th>
<th>Number of individuals</th>
<th>mean size</th>
<th>Percentage female Headed Household</th>
<th>Single person HH</th>
<th>HHH mean years education</th>
<th>Dependency ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fieldwork</strong></td>
<td>52</td>
<td>573</td>
<td>11.23</td>
<td>27.5%</td>
<td>2%</td>
<td>6.67</td>
<td>1.11</td>
</tr>
<tr>
<td><strong>DHS definition</strong></td>
<td>104</td>
<td>490</td>
<td><strong>5.86</strong></td>
<td>41.9%</td>
<td>23.1%</td>
<td>7.17</td>
<td>1.20</td>
</tr>
<tr>
<td><strong>Census definition</strong></td>
<td>133</td>
<td>421*</td>
<td>5.64</td>
<td>46.3%</td>
<td>27%</td>
<td>7.18</td>
<td>1.27</td>
</tr>
</tbody>
</table>

*152 would have been captured in other households
Modelling scenarios

- The Tanzanian statistical definition of household reduces the average household size
- Increases the proportion of female headed HHs
- Distorts the characteristics of household heads
- Disassociates people from resources to which they have access
- Often single person’s household linked to bigger more complex structure
2. DHS data: Thinking creatively

• Compare and contrast indicators at de jure and de facto level: how do they impact the outcome?

• Analyse specific subgroups:
  – Single persons households
  – Polygamous unions
  – De jure members that did not sleep in the household the night before
  – Female headed households

• Objective twofold: exploit existing data and understand how subsamples characteristics might bias the outcome
How do different samples affect the gender composition of the household?

**HH gender composition**

- **Male**
  - Single person HH: 60
  - Did not sleep there sample: 50
  - Overall sample: 40
  - Female Headed HH: 30

- **Female**
  - Male: 50
  - Did not sleep there sample: 40
  - Overall sample: 30
  - Female Headed HH: 20
Where do they live?

Place of residence

- Single person HH
- Did not sleep there last night
- Overall sample
- Female headed HH

Bar chart showing the distribution of place of residence for different groups (Single person HH, Did not sleep there last night, Overall sample, Female headed HH) across Capital, large city, Small city, Town, and Countryside.
How wealthy are they?

Wealth quintile

- Single person HH
- Did not sleep here last night
- Overall sample
- Female headed HH
# Results DHS data analysis-scenarios

<table>
<thead>
<tr>
<th></th>
<th>% included</th>
<th>mean age sample</th>
<th>Dependency ratio</th>
<th>Mean years of HH education</th>
<th>% female population</th>
<th>Age head of HH</th>
<th>house hold size</th>
<th># people per room</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Sample</td>
<td></td>
<td>25.05</td>
<td>1.26</td>
<td>4.88</td>
<td>51</td>
<td>44.73</td>
<td>5.13</td>
<td>2.48</td>
</tr>
<tr>
<td>de jure</td>
<td>96.3</td>
<td>25.09</td>
<td>1.35</td>
<td>-</td>
<td>50.8</td>
<td>44.68</td>
<td>4.94</td>
<td>2.3</td>
</tr>
<tr>
<td>de facto</td>
<td>93.2</td>
<td>21.74</td>
<td>1.55</td>
<td>4.33</td>
<td>51.8</td>
<td>39.99</td>
<td>4.78</td>
<td>2.4</td>
</tr>
<tr>
<td>Female HHH</td>
<td>24.2</td>
<td>23.37</td>
<td>1.47</td>
<td>3.01</td>
<td>61.2</td>
<td>45.59</td>
<td>4.58</td>
<td>2.42</td>
</tr>
</tbody>
</table>
Light at the end of the tunnel?
Ways of dealing with ‘fuzzy’ household at the collection stage

• Collect information on who resides in the household as reported by the respondent before being selected for the main part of the questionnaire
  • The DHS, for example, uses the households to select the individuals. The first part could be expanded to include more information
Ways of dealing with ‘fuzzy’ household at the collection stage

• Collect data in more sensible way that allows better configurations
  – include information on who slept there the night before, who ate and possibly on contributions to the household economy
  – Relationship to hh head
  – Line numbers and relationship to each other

• Where possible and in particular for specialized surveys avoid assumptions of crisp boundaries – allow multiple membership of HHs and find ways to record it (e.g: Hosegood & Timaeus).
Ways of dealing with ‘fuzzy’ household at the analysis stage

• Education of users: more background material on the issues surrounding the impact of the household definition
  – Careful interpretation of the results
  – Non-technical language to educate policy makers on the interpretation of the data
Ways of dealing with ‘fuzzy’ household at the analysis stage

• Methodological material available to users
  – Warnings from users’ manuals
  – Make better use of the household recode of the DHS survey when analysing individual files
  – More methodological research into the use of households needed
  – There is a limited literature on the impact of the definition on the possible outcomes. Especially poverty mapping
  – Future research needed into how different types of respondents can influence the household’s composition structure (e.g.: example of man not reporting wife’s son).
Discussion and few thoughts

• **NOT** trying to redefine the household
• More awareness on the issues needed
  – Flexible thinking
• More methodological developments needed
  – Flexible collection

‘The household is central to the development process. Not only is the household a production unit but it is also a consumption, social and demographic unit’ Kenya: Ministry of Planning and National Development 2003, p59