Use of CAPI for agricultural surveys

Is CAPI right for my survey?
Ground rules:

Training is a big investment. To learn and benefit we agree to:

1. **Ask questions** if we don’t understand
2. **Take a break** if we aren’t concentrating
3. Respect our colleagues – **no phones. Be on time.**
4. **Participate** – be in the room
5. **Let us know** if anything is not working
6. Have some **fun**
Outline

• Why some still use PAPI
• Benefits of CAPI
• Can CAPI reduce costs?
• CAPI technical issues
• Survey Solutions and brief comparisons to other products
Why some still use PAPI

• Small, simple surveys
• Little or no Internet or cell phone accessibility
• Avoid the initial cost of tablets
• Scheduling issues
• Reliance on existing methods and skills
Benefits of CAPI

• Can provide immediate checks on routing, ranges, and consistency during the interview
• Can provide faster and more consistent field editing, and
• Can provide faster and more consistent home office (aggregated data) editing, so
• Should provide higher data quality
• Faster availability of final data
• Minimum paper questionnaires to manage as back-up:
  — paper questionnaire can be used in emergency:
    • Frozen devices (tablets, pda, smartphone...)
    • power issues (dead battery, no power source around...)
    • Etc.
Some Additional Benefits of CAPI

(Depending on software and hardware)

• Collect sensor data directly
  • Weight
  • GPS (including land area)
  • Date/time
  • Soil quality
  • Water quality
  • Temperature
  • Others...

• Use media (photos, videos, sound) as part of the interview
Can CAPI reduce costs?

- No double (verification) data entry and corrections
- Far lower paper, printing, and questionnaire storage costs
- Shorter interview durations can reduce fieldwork costs
- Lower logistics costs: no transporting and storing paper questionnaires
- Reduced waiting period for production of final data
Can CAPI reduce costs?

Cost comparisons of PAPI versus CAPI survey data collection (in 2011 purchasing power parity (PPP) US$)

<table>
<thead>
<tr>
<th>Description</th>
<th>Wave 1</th>
<th>Wave 2</th>
<th>Wave 3</th>
<th>Wave 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>2009-2010</td>
<td>2010-2011</td>
<td>2012-2013</td>
<td>2013-2014</td>
</tr>
<tr>
<td>Sample Size</td>
<td>3000</td>
<td>3000</td>
<td>3000</td>
<td>3000</td>
</tr>
<tr>
<td>Number of questions</td>
<td>1090</td>
<td>1097</td>
<td>1152</td>
<td>1148</td>
</tr>
<tr>
<td>Method of data collection</td>
<td>PAPI</td>
<td>CAPI</td>
<td>CAPI</td>
<td>CAPI</td>
</tr>
<tr>
<td>Costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enumerators salaries</td>
<td>91.46</td>
<td>81.05</td>
<td>63.96</td>
<td>61.42</td>
</tr>
<tr>
<td>Supervisor salaries</td>
<td>45.27</td>
<td>40.12</td>
<td>31.66</td>
<td>30.40</td>
</tr>
<tr>
<td>Data Entry</td>
<td>28.38</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cleaning costs</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Paper questionnaire cost</td>
<td>35.04</td>
<td>6.21</td>
<td>4.44</td>
<td>3.68</td>
</tr>
<tr>
<td>Electronic equipment cost</td>
<td>-</td>
<td>8.14</td>
<td>7.26</td>
<td>7.07</td>
</tr>
<tr>
<td>Cost per interview</td>
<td>200.16</td>
<td>135.52</td>
<td>107.33</td>
<td>102.56</td>
</tr>
</tbody>
</table>

Note: Values converted from Ugandan Shillings to 2005 PPP values using the World Development Indicators (2016).
Costs specific to CAPI

• Initial investment in tablets, server, and service provider accounts
• Increased preparation time for connectivity, tablet setup, server setup, server maintenance
• New system for many organizations to learn
  • Different set of management skills
  • Depending on software, can require additional programming and other technical skills
• Increased need for interviewer and field supervisor training
• Overall cost depends on learning curve, software cost, and preparation
Costs specific to CAPI

• Direct Costs
  – Tablets (purchase, import, and setup)
  – Software
  – Training
  – Communications and connectivity
  – Server acquisition, setup, and maintenance

• Indirect Costs
  – Technical Support
CAPI Technical Issues

• Cell phone or Internet access
• Tablet availability
• Server availability and maintenance
• Availability of capable interviewers and office staff
Survey Solutions

• Very easy to learn to use, little programming or technical assistance required
• Programming uses C# syntax but is generally intuitive
• Enables complex skips (routing)
• Enables complex consistency checks
Survey Solutions

• Includes built-in case management system that is easy to configure and use
  – Sample definition
  – Case assignments and tracking
  – Case approvals and rejections
Survey Solutions

• Includes built-in case management system that is easy to configure and use
  – Report generation
    • Survey progress tracking
    • Durations of interviews
    • Locations of interviews displaying maps
Survey Solutions

• Server can provide extensive “paradata,” or data about the data
  – Can be exceptionally valuable for quality control, and includes:
    • Time stamps for each question
    • Interview durations
    • Question sequences
  – Enables creation of QC reports including:
    • Interview duration
    • Numbers of invalid responses (even if corrected)
    • By time of day, interviewer, team, or location using maps
Comparisons with Survey Solutions

- **Survey Solutions**
  - Produced and supported by the World Bank

- **CSPro**
  - Produced and supported by the US Bureau of Census

- **ODK**
  - Open source program developed at the University of Washington

- **Blaise**
  - Produced and supported by Statistics Netherlands
## Comparisons with Survey Solutions

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Survey Solutions</th>
<th>CSPro</th>
<th>ODK</th>
<th>Blaise</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Software license</strong></td>
<td>Free</td>
<td>Free</td>
<td>Free</td>
<td>$$ $$</td>
</tr>
<tr>
<td><strong>Server requirements</strong></td>
<td>HQ and Supervisor server can be anywhere</td>
<td>FTP, Dropbox</td>
<td>ODK Aggregate server</td>
<td>IIS, ASP, relational database server</td>
</tr>
<tr>
<td><strong>Learning curve</strong></td>
<td>Shallow, uses simplified C# syntax</td>
<td>Steep, uses internal language</td>
<td>Moderate, uses MS Excel</td>
<td>Steep, uses internal language</td>
</tr>
<tr>
<td><strong>Development time</strong></td>
<td>Shortest</td>
<td>Long</td>
<td>Relatively short</td>
<td>Long</td>
</tr>
<tr>
<td><strong>Android questionnaire testing</strong></td>
<td>Included</td>
<td>Not included</td>
<td>Not included</td>
<td>Windows only</td>
</tr>
<tr>
<td><strong>Questionnaire programming skills required</strong></td>
<td>Minimal training required, may use well-known C#</td>
<td>Training required, uses unique language</td>
<td>Uses Excel, some training needed</td>
<td>Training required, uses unique language</td>
</tr>
<tr>
<td><strong>Case management</strong></td>
<td>Built in</td>
<td>Use CSPro program</td>
<td>External program</td>
<td>Use Blaise program</td>
</tr>
<tr>
<td><strong>Export data formats</strong></td>
<td>SPSS, Stata, TXT</td>
<td>SPSS, SAS, Stata, XML, CSV R, CSV, TXT</td>
<td>Text, XML, ASCIIIRelational, OLEDB</td>
<td></td>
</tr>
</tbody>
</table>
Questions?