Living Standards Measurement Study-Integrated Surveys on Agriculture: Innovations Built on Tradition

Innovations In Survey Design for Policy
PREM week
April 26, 2011
MOTIVATION

Importance of agriculture for poverty reduction, yet:

- Poor data: low quality, inadequate periodicity and comparability, lack of policy relevance
- Failure to address high levels of diversification, linkages to non-farm, poverty, health, …
- Lack of panel data
- Institutional constraints in agricultural data production and analysis
- Lack of analytical capacity
- Poor dissemination of data and findings
- Overall, too little attention to agriculture and agricultural statistics
OPTIONS

- Strengthen Farm Surveys
  - Agricultural production units (rural) only
  - Single sector coverage
  - Weak statistical capacity
  - Organizational structure/data quality control

- Strengthen Multi-purpose Surveys (LSMS)
  - Thematic trade-offs
  - Sampling issues (minor crops, commercial farms, …)
  - Global Strategy for the Improvement of Agricultural and Rural Statistics
    - Integration

- LSMS-Integrated Surveys on Agriculture
APPROACH

- In the tradition of the LSMS, not “one-size-fits-all”
- Build on existing systems
- Cross-country “convergence” over time
- Inter-institutional, inter-agency effort
- Work on 4 fronts:
  - Household survey data collection
  - Methodological validation/research
  - Capacity Building
  - Dissemination
MAIN FEATURES

• Panel
  – Frequency (varies by country)
  – Tracking of movers
  – Tracking of split-offs

• Sample design
  – Population-based frame
  – Sample size
    • Relatively small at baseline (3,000-5,000 HHs)
    • Few domains of inference
MAIN FEATURES (cont’d)

- Integrated approach
  - Multi-topic survey instrument
    - Agriculture plus non-farm, poverty, nutrition, *inter alia*
  - Build on existing/planned surveys
    - Local partners: National Statistical Office, MoA
    - National Strategy for the Development of Statistics (NSDS)
  - Improved links to other data sources
    - PopCensus, AgCensus (Small Area Estimation)
    - Geo-referencing
    - Multiple instruments
MAIN FEATURES (cont’d)

- From Centralized to Field-based data entry (concurrent) to Computer Assisted Personal Interviews (CAPI)
- Use of GPS for plot location and measurement
- Public Access Data Policy
- Inter-agency collaborations
  - Global: WFP, IFAD, WFC, IFPRI, FAO, ARD, …
  - In-country: USAID, DFID, GTZ, Stat Norway, …
COUNTRIES

- Tanzania National Panel Survey: wave 2 in the field; wave 1 data available
- Uganda National Panel Survey: Wave 2 in the field, wave 1 linked to 2005/06 UNHS
- Malawi Integrated Panel Household Survey: Wave 1 completed
- Nigeria General Household Survey Panel: Wave 1 completed
- Ethiopia Agriculture Sample Survey: Wave 1: Fall 2011
- Mali Integrated Agricultural Survey: Spring 2012; pilot activities in Fall 2011
METHODOLOGY

• Recall vs. Diary (vs. crop cutting) for production estimates
• GPS vs. self-reported area measurement (vs. compass and rope)
• Use of mobile phones
• Measurement of income components
• Livestock
  – Milk
  – Stocks by breed
  – Pastoralists
SOURCEBOOKS

- Tracking
- Sampling Weights in Panel Surveys
- Fishery
- Livestock
- Climate Change
  - “Land-based”
  - “Water-based”
Example 1: USING GPS

- Non-sampling errors in agricultural statistics known to be large
  - data collection methodology and techniques
  - interviewer’s effect, respondents’ interpretation questions
  - motivation/incentive to provide accurate answers
- Use of GPS offers opportunity to improve land, and thus yield, measures
- Differences with self-reporting may be substantial
- Difference varies by farm size, thus potential effect on yields and Inverse Farmsize-Productivity Relationship (IR)
## GPS vs. self-reported area measurements

<table>
<thead>
<tr>
<th>Deciles</th>
<th>Area of HH landholding</th>
<th>Nb. of plots per hh</th>
<th>Mean farm area Using GPS</th>
<th>Mean farm area using Self-Reported</th>
<th>Farm Discrepancy (-Self Reported)</th>
<th>Discrepancy in % terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.01-0.65</td>
<td>1.70</td>
<td>0.37</td>
<td>0.73</td>
<td>-0.36</td>
<td>-97%</td>
</tr>
<tr>
<td>2</td>
<td>0.66-1.12</td>
<td>2.33</td>
<td>0.90</td>
<td>1.43</td>
<td>-0.53</td>
<td>-59%</td>
</tr>
<tr>
<td>3</td>
<td>1.13-1.62</td>
<td>2.40</td>
<td>1.37</td>
<td>1.78</td>
<td>-0.41</td>
<td>-30%</td>
</tr>
<tr>
<td>4</td>
<td>1.63-2.09</td>
<td>2.70</td>
<td>1.84</td>
<td>2.36</td>
<td>-0.52</td>
<td>-28%</td>
</tr>
<tr>
<td>5</td>
<td>2.09-2.69</td>
<td>2.94</td>
<td>2.38</td>
<td>2.91</td>
<td>-0.52</td>
<td>-22%</td>
</tr>
<tr>
<td>6</td>
<td>2.7</td>
<td>2.80</td>
<td>3.04</td>
<td>3.53</td>
<td>-0.48</td>
<td>-16%</td>
</tr>
<tr>
<td>7</td>
<td>3.44-4.57</td>
<td>2.74</td>
<td>3.96</td>
<td>4.10</td>
<td>-0.14</td>
<td>-4%</td>
</tr>
<tr>
<td>8</td>
<td>4.59-6.16</td>
<td>2.94</td>
<td>5.31</td>
<td>5.18</td>
<td>0.13</td>
<td>2%</td>
</tr>
<tr>
<td>9</td>
<td>6.17-9.13</td>
<td>3.20</td>
<td>7.46</td>
<td>7.08</td>
<td>0.39</td>
<td>5%</td>
</tr>
<tr>
<td>10</td>
<td>9.14-600</td>
<td>3.40</td>
<td>21.03</td>
<td>17.07</td>
<td>3.96</td>
<td>19%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>0.01-600</strong></td>
<td><strong>2.70</strong></td>
<td><strong>4.75</strong></td>
<td><strong>4.60</strong></td>
<td><strong>0.15</strong></td>
<td><strong>3%</strong></td>
</tr>
</tbody>
</table>
## Yield and farmsize

<table>
<thead>
<tr>
<th>Landholding</th>
<th>Average land areas</th>
<th>Yield (GPS)</th>
<th>Yield (Self Reported)</th>
<th>Bias in yield (GPS-Self Reported)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Farms</td>
<td>0.01-1.45</td>
<td>0.7</td>
<td>236</td>
<td>170</td>
</tr>
<tr>
<td>Medium Farms</td>
<td>1.46-3.57</td>
<td>2.4</td>
<td>208</td>
<td>193</td>
</tr>
<tr>
<td>Large Farms</td>
<td>3.58-600</td>
<td>10.3</td>
<td>77</td>
<td>100</td>
</tr>
</tbody>
</table>
Bias in land measurement: “Heaping”

Plot Size Measured with GPS and Farmers’ Estimate

- **GPS**
- **Farmers’ Estimate**

% (Area self Reported)

<table>
<thead>
<tr>
<th>% (Area GPS)</th>
<th>% (Area self Reported)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

Plot showing the comparison between GPS measured area and farmers’ estimated area with a histogram and line graph.
Testing the IR

\[
\ln \frac{Y_i}{A_i} = \beta_0 + \beta_1 \ln A_i + \beta_2 X_i + \beta_3 B + u_i
\]

\( \frac{Y_i}{A_i} \) = Yield (value of output per acre)  
\( A_i \) = land area cultivated  
\( X_i \) = vector of HH characteristics  
\( B_i \) = Bias land  
\( U_i \) = error term
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dep. Variable Agric. Yield</strong></td>
<td>Self-Reported</td>
<td>GPS</td>
</tr>
<tr>
<td>Log Land Size</td>
<td>-0.62***</td>
<td>-0.83***</td>
</tr>
<tr>
<td>Observations</td>
<td>2860</td>
<td>2861</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.60</td>
<td>0.63</td>
</tr>
</tbody>
</table>
Yields and farm size

UGANDA: Inverse Farm Size Productivity Relationship

- Yield vs. Deciles of Land Cultivated
- Land Self-Reported vs. Land GPS
Example 2: USING DIARIES

• Long recall period
• Recall particularly difficult for continuous and repeated crops (tubers, vegetables, …)
• Take advantage of NPS multiple visits
• Crop card monitors
  – Weekly visit
• Non-standard units of measurement
### Diaries vs. recall

<table>
<thead>
<tr>
<th></th>
<th>Frequency (%)</th>
<th>Production value ($)</th>
<th>No. of entries</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Diary</td>
<td>Recall</td>
<td>Diary</td>
</tr>
<tr>
<td><strong>Cash crops</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coffee</td>
<td>21.69</td>
<td>31.89</td>
<td>10.63</td>
</tr>
<tr>
<td>Rice</td>
<td>12.29</td>
<td>5.43</td>
<td>15.74</td>
</tr>
<tr>
<td>Cotton</td>
<td>0.57</td>
<td>9.14</td>
<td>0.5</td>
</tr>
<tr>
<td>Sugarcane</td>
<td>17.36</td>
<td>4.21</td>
<td>0.37</td>
</tr>
<tr>
<td>All cash crops</td>
<td></td>
<td></td>
<td>27.24</td>
</tr>
<tr>
<td><strong>Food crops seasonal</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maize</td>
<td>73.93</td>
<td>76.8</td>
<td>66.45</td>
</tr>
<tr>
<td>Groundnuts</td>
<td>41.42</td>
<td>30.37</td>
<td>46.94</td>
</tr>
<tr>
<td>Beans</td>
<td>78.36</td>
<td>71.67</td>
<td>124.38</td>
</tr>
<tr>
<td>All seasonal food</td>
<td></td>
<td></td>
<td>297.43</td>
</tr>
<tr>
<td><strong>Food crop continuous</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Banana</td>
<td>75.62</td>
<td>59.4</td>
<td>120.03</td>
</tr>
<tr>
<td>Sweet potatoes</td>
<td>83.16</td>
<td>59.02</td>
<td>81.09</td>
</tr>
<tr>
<td>Cassava</td>
<td>82.01</td>
<td>58.82</td>
<td>55.36</td>
</tr>
<tr>
<td>All cont. food</td>
<td></td>
<td></td>
<td>275.58</td>
</tr>
<tr>
<td><strong>Fruit &amp; Vegetables</strong></td>
<td></td>
<td></td>
<td>14.87</td>
</tr>
<tr>
<td><strong>All crops</strong></td>
<td></td>
<td></td>
<td>615.12</td>
</tr>
</tbody>
</table>
FOCUS OF FUTURE WORK

• Soil fertility measurements
• GPS measurements
  – Bias in measurement
  – Effect of slope and weather
  – Small plots
• Labor inputs
• Water measurements
• Cognitive and non-cognitive skills
CHALLENGES

• Integration
• Institutional framework
• Analytical capacity
• Level of representativeness (sampling)
• Burden on respondents
• Donor coordination
• Managing expectations