Using High Frequency Data to Measure Resilience
ICT4D, Hyderabad

Erwin Knippenberg ¹

¹Phd Candidate
Cornell University

May 2017
Outline

Why it Matters

MIRA project

Dashboard

Prediction Using Machine Learning

Conclusion
Why it Matters

Figure: Drought Projections by end of century (Dai 2011)
What is Resilience?

Figure: Stylized recovery trajectory

Source: Hoddinott (2014a)
Notes: FCS = food security score; HH-Q = household Q; HH-R = household R.
Data requirements

- Pre and Post Shock
Data requirements

- Pre and Post Shock
- Measure of shock severity
Data requirements

- Pre and Post Shock
- Measure of shock severity
  - Objective (geo-spatial)
  - Subjective (reported)
Data requirements

- Pre and Post Shock
- Measure of shock severity
  - Objective (geo-spatial)
  - Subjective (reported)
- Measure of interest (income, damage to structure etc..)
Data requirements

- Pre and Post Shock
- Measure of shock severity
  - Objective (geo-spatial)
  - Subjective (reported)
- Measure of interest (income, damage to structure etc..)
- Intervention or Resilience Capacities
Illustration: Ethiopia

Figure: Effect of Cash Transfer on Resilience
The MIRA project

Contributions

- **High frequency** data collection using local enumerators with smart-phones
The MIRA project

Contributions

- High frequency data collection using local enumerators with smart-phones
- Up-to-date Dashboard, disseminated to communities
The MIRA project
Contributions

- **High frequency** data collection using local enumerators with smart-phones
- Up-to-date **Dashboard**, disseminated to communities
- Predict future drought using **Machine Learning**
Context: Chikwawa, Malawi
Data Collection

Figure: MIRA enumerator using CommCare
Shocks Experienced

Figure: Evolution of Shock Incidence over time
### Dashboard Dissemination

**Figure:** Snapshot of MIRA Dashboard

<table>
<thead>
<tr>
<th>Last Updated</th>
<th>5/15/17</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>District</th>
<th>crop_disease</th>
<th>drought</th>
<th>end_assistar flood</th>
<th>hh_breakup</th>
<th>hh_death</th>
<th>illness</th>
<th>livestock_die</th>
<th>rise_food_prices</th>
<th>theft</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chikwawa</td>
<td>22%</td>
<td>75%</td>
<td>4%</td>
<td>23%</td>
<td>8%</td>
<td>27%</td>
<td>7%</td>
<td>33%</td>
<td>4%</td>
</tr>
<tr>
<td>TA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ngabu</td>
<td>28%</td>
<td>74%</td>
<td>2%</td>
<td>14%</td>
<td>1%</td>
<td>3%</td>
<td>26%</td>
<td>5%</td>
<td>22%</td>
</tr>
<tr>
<td>Lundu</td>
<td>12%</td>
<td>85%</td>
<td>3%</td>
<td>43%</td>
<td>2%</td>
<td>18%</td>
<td>22%</td>
<td>3%</td>
<td>21%</td>
</tr>
<tr>
<td>Makuwira</td>
<td>25%</td>
<td>65%</td>
<td>7%</td>
<td>17%</td>
<td>2%</td>
<td>6%</td>
<td>33%</td>
<td>9%</td>
<td>54%</td>
</tr>
<tr>
<td>Maseya</td>
<td>24%</td>
<td>100%</td>
<td>5%</td>
<td>5%</td>
<td>10%</td>
<td>0%</td>
<td>33%</td>
<td>29%</td>
<td>29%</td>
</tr>
<tr>
<td>GVH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kalulu</td>
<td>21%</td>
<td>74%</td>
<td>3%</td>
<td>33%</td>
<td>0%</td>
<td>5%</td>
<td>23%</td>
<td>8%</td>
<td>18%</td>
</tr>
<tr>
<td>Nyambalo</td>
<td>3%</td>
<td>72%</td>
<td>3%</td>
<td>5%</td>
<td>0%</td>
<td>5%</td>
<td>44%</td>
<td>0%</td>
<td>97%</td>
</tr>
<tr>
<td>Champhanda</td>
<td>36%</td>
<td>80%</td>
<td>12%</td>
<td>4%</td>
<td>4%</td>
<td>0%</td>
<td>20%</td>
<td>8%</td>
<td>24%</td>
</tr>
<tr>
<td>Mpama</td>
<td>40%</td>
<td>38%</td>
<td>3%</td>
<td>18%</td>
<td>8%</td>
<td>0%</td>
<td>30%</td>
<td>18%</td>
<td>43%</td>
</tr>
<tr>
<td>Jombo</td>
<td>68%</td>
<td>86%</td>
<td>0%</td>
<td>0%</td>
<td>2%</td>
<td>4%</td>
<td>26%</td>
<td>4%</td>
<td>22%</td>
</tr>
<tr>
<td>Bestala</td>
<td>42%</td>
<td>84%</td>
<td>5%</td>
<td>92%</td>
<td>3%</td>
<td>18%</td>
<td>16%</td>
<td>5%</td>
<td>26%</td>
</tr>
<tr>
<td>Chagambo</td>
<td>21%</td>
<td>92%</td>
<td>5%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>5%</td>
<td>3%</td>
<td>37%</td>
</tr>
<tr>
<td>Kanyimbi</td>
<td>3%</td>
<td>78%</td>
<td>0%</td>
<td>25%</td>
<td>0%</td>
<td>6%</td>
<td>38%</td>
<td>3%</td>
<td>34%</td>
</tr>
<tr>
<td>Sabala</td>
<td>53%</td>
<td>38%</td>
<td>24%</td>
<td>50%</td>
<td>3%</td>
<td>24%</td>
<td>62%</td>
<td>24%</td>
<td>79%</td>
</tr>
<tr>
<td>Malikopo</td>
<td>10%</td>
<td>69%</td>
<td>3%</td>
<td>0%</td>
<td>0%</td>
<td>3%</td>
<td>28%</td>
<td>0%</td>
<td>26%</td>
</tr>
<tr>
<td>Biyasi</td>
<td>14%</td>
<td>90%</td>
<td>0%</td>
<td>29%</td>
<td>10%</td>
<td>5%</td>
<td>52%</td>
<td>5%</td>
<td>33%</td>
</tr>
<tr>
<td>Bilati</td>
<td>0%</td>
<td>85%</td>
<td>3%</td>
<td>85%</td>
<td>0%</td>
<td>0%</td>
<td>13%</td>
<td>0%</td>
<td>13%</td>
</tr>
<tr>
<td>Chapamoko</td>
<td>17%</td>
<td>61%</td>
<td>3%</td>
<td>0%</td>
<td>3%</td>
<td>0%</td>
<td>33%</td>
<td>11%</td>
<td>28%</td>
</tr>
<tr>
<td>Mafale</td>
<td>0%</td>
<td>87%</td>
<td>3%</td>
<td>1%</td>
<td>0%</td>
<td>33%</td>
<td>22%</td>
<td>0%</td>
<td>21%</td>
</tr>
<tr>
<td>M’bande</td>
<td>24%</td>
<td>100%</td>
<td>5%</td>
<td>5%</td>
<td>10%</td>
<td>0%</td>
<td>33%</td>
<td>29%</td>
<td>29%</td>
</tr>
<tr>
<td>Sekeni</td>
<td>20%</td>
<td>67%</td>
<td>0%</td>
<td>40%</td>
<td>0%</td>
<td>7%</td>
<td>13%</td>
<td>20%</td>
<td>7%</td>
</tr>
</tbody>
</table>
Dashboard

Figure: Disseminating Dashboard to Communities
Predicting Future Shock Incidence

- In addition to responding, we want to **anticipate**
Predicting Future Shock Incidence

- In addition to responding, we want to **anticipate**
- Use previous rounds to predict shock incidence in future
Predicting Future Shock Incidence

- In addition to responding, we want to anticipate
- Use previous rounds to predict shock incidence in future
- Horse-race between three algorithms:
Predicting Future Shock Incidence

- In addition to responding, we want to **anticipate**
- Use previous rounds to predict shock incidence in future
- Horse-race between three algorithms:
  - K Nearest Neighbor
Predicting Future Shock Incidence

- In addition to responding, we want to **anticipate**
- Use previous rounds to predict shock incidence in future
- Horse-race between three algorithms:
  - K Nearest Neighbor
  - Naive Bayes
Predicting Future Shock Incidence

- In addition to responding, we want to **anticipate**
- Use previous rounds to predict shock incidence in future
- Horse-race between three algorithms:
  - K Nearest Neighbor
  - Naive Bayes
  - LASSO
Predicted vs. Actual Incidence of Drought

Figure: Actual vs. Predicted Incidence of Drought in Chikwawa
Predicted vs. Actual Incidence of Drought

Figure: Actual vs. Predicted Incidence of Drought in Chikwawa
Conclusion

**Three contributions:**

1. High Frequency data collection using local enumerators with smart-phones
2. Real time upload to the cloud permits rapid feedback
Thank You!