Heterogeneity

Hard to think that agricultural technologies have the same returns across any given population of farmers

We know that average returns to agricultural technologies such as hybrid maize and fertilizer can be extremely high, on the order of 100%

If you are a farmer, using these technologies can double your profits from maize!

Is that true for all farmers?
How About These Farmers?

And These Farmers?
And These?

Understanding Heterogeneity

Can heterogeneity in the benefits (and costs) of technologies explain one of the main puzzles that we see all over Africa, i.e. that

i. Despite high average returns, a significant number of households do not use these technologies: why not?

ii. Adoption rates show no acceleration/increases over time

The answer turns out to be yes!
Outline

Motivating the puzzle and how heterogeneity can account for this puzzle

Model of technology adoption with heterogeneity

Estimating the model

Correlate the estimated returns w/ adoption: does heterogeneity in costs and benefits explain adoption and lack of adoption?

What is My Hypothesis?

Are there differences in returns to the technology across farmers?

If so, is it simply the case that the farmers who do not use it simply do not benefit from it?

Is the lack of adoption explained by zero or negative returns to the technology?
Why is This Important?

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Kenya</td>
<td>Maize</td>
<td>0.362</td>
<td>2.373</td>
<td>1.169</td>
<td>-1.198</td>
</tr>
<tr>
<td>India</td>
<td>Wheat</td>
<td>4.876</td>
<td>2.514</td>
<td>3.343</td>
<td>1.235</td>
</tr>
<tr>
<td>India</td>
<td>Rice</td>
<td>0.954</td>
<td>1.714</td>
<td>3.310</td>
<td>0.838</td>
</tr>
<tr>
<td>Mexico</td>
<td>Maize</td>
<td>2.057</td>
<td>4.267</td>
<td>-0.548</td>
<td>1.447</td>
</tr>
<tr>
<td>Zambia</td>
<td>Maize</td>
<td>-0.267</td>
<td>10.40</td>
<td>1.571</td>
<td>-1.707</td>
</tr>
</tbody>
</table>

Hybrid Adoption

![Graph showing Hybrid Adoption from 1995 to 2005 for different regions such as Rift Valley, Western, Central, Sample, Eastern, and Nyanza. The graph plots the fraction of seed that is hybrid over the years.](image-url)
What About the Farmers?

<table>
<thead>
<tr>
<th>Priorities to Improve Family Well Being</th>
<th>HHs for whom it is the Top Priority</th>
<th>HHs that Place it in the Top 3 Priorities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase Yields on Existing Land</td>
<td>39.4%</td>
<td>72.7%</td>
</tr>
<tr>
<td>Obtain More Land</td>
<td>16.4%</td>
<td>29.2%</td>
</tr>
<tr>
<td>Obtain More Animals</td>
<td>14.5%</td>
<td>55.4%</td>
</tr>
<tr>
<td>Start a Business/ Earn more from Business</td>
<td>23.2%</td>
<td>48.4%</td>
</tr>
<tr>
<td>Education</td>
<td>2.3%</td>
<td>5.9%</td>
</tr>
<tr>
<td>Credit</td>
<td>0.3%</td>
<td>0.5%</td>
</tr>
</tbody>
</table>
Yields by Sector, 2004

How Do I Test My Hypothesis
What Do I Do?

I start with an economic model of how households make decisions about which technology to use.

Adoption Decisions I

Model is based on a generalized Roy model of sector (technology) choice.

Allows for two forms of household specific heterogeneity:
- Absolute advantage (not correlated with the technology choice)
- Comparative advantage (relative productivity in hybrid over non-hybrid)
Adoption Decisions II

A farmer will adopt hybrid if his profits from hybrid are greater than his profits from non hybrid

Allow directly for the following:
- Different quantities of fertilizer to be used
- Fixed (over time) costs of access
- Time varying costs of access

Prices of the two seeds are different, but not much spatial variation in this price difference

Adoption Decisions III

This is basically a model of comparative advantage that finally gives me:

A farmer will adopt if \[ \psi \theta_i > A_{it} + \Delta_{it} - (\beta^H_i - \beta^N_i) \]

The parameter \( \psi \) tells us ...
Adoption Decisions IV

Implies a production function for maize of the following form:

\[ y_{it} = \beta_t^N + \alpha_i + (\beta_t^H - \beta_t^N)h_{it} + X_{it}' \gamma^N + \phi \theta_i h_{it} + X_{it}'(\gamma^H - \gamma^N)h_{it} + \varepsilon_{it} \]

Interpretation

\[ y_{it} = \pi + \beta h_{it} + \theta_i + \tau_i + \phi \theta_i h_{it} + \varepsilon_{it} \]

\[ = \alpha_i, \text{ HH Specific Intercept (HH specific average yield)} \]

\[ \text{HH Specific Slope (HH specific gain to hybrid)} \]

The covariance between the household specific slopes and intercepts:

\[ \text{cov}(\alpha_i, \phi \theta_i) = \phi \sigma^2 \]

Structural coefficient \( \phi \) tells us if high intercept HHs are high slope HHs.
What Regressions to Run

Sample Villages
Transitions in Hybrid (Switching)

<table>
<thead>
<tr>
<th>Fraction of Sample (%)</th>
<th>Transition</th>
</tr>
</thead>
<tbody>
<tr>
<td>20.38</td>
<td>000</td>
</tr>
<tr>
<td>2.83</td>
<td>001</td>
</tr>
<tr>
<td>6.07</td>
<td>010</td>
</tr>
<tr>
<td>5.99</td>
<td>100</td>
</tr>
<tr>
<td>4.91</td>
<td>011</td>
</tr>
<tr>
<td>3.16</td>
<td>101</td>
</tr>
<tr>
<td>7.15</td>
<td>110</td>
</tr>
<tr>
<td>49.50</td>
<td>111</td>
</tr>
</tbody>
</table>

What Regressions to Run?

- Regression of output/acre on hybrid use: coefficient of 1.024
- Control for inputs (labor, fertilizer, etc.): coefficient of 0.556
- Control for a household “fixed effect”: coefficient of 0.149
- Instrumental variables: coefficient of 1.452
- Many are significantly different from each other, but none analyzes what the return to hybrid is if it differs across farmers
Estimation (Generalize Chamberlain)

The basic yield function for the heterogeneous returns model is:

\[ y_{it} = \delta + \theta_i + \beta h_{it} + \phi \theta_i h_{it} + \tau_i + \epsilon_{it} \]

Project \( \theta_i \) onto the history and the interactions of the hybrid histories:

\[ \theta_i = \lambda_0 + \lambda_1 h_{i1} + \lambda_2 h_{i2} + \lambda_3 h_{i1} h_{i2} + \nu_i \]

Use Minimum Distance:
- Run a reduced form for each time period (includes all hybrid histories) to get six reduced form parameters
- These map onto five structural parameters (\( \phi \beta \lambda_1 \lambda_2 \lambda_3 \))
- The structural parameters are over-identified

What Do I Find?

YES! Dramatic differences in what returns to hybrid would be

Would those not using the technology have zero returns? NO!

Some of them do! The rest would have even higher than average returns from using hybrid – puzzle deepens!

These farmers have high costs that are not reflected in prices (poor access to the technology)

Points to these farmers being pretty “rational”!
Comparative Advantage I

Implied Returns \((\beta + \varphi \theta_i)\)
Comparative Advantage II

What Costs or Constraints?

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>H</th>
<th>L</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dist to Fertilizer (x100)</td>
<td>-0.52 (0.081)</td>
<td>-0.51 (0.081)</td>
<td>-0.52 (0.081)</td>
<td>-0.497 (0.082)</td>
</tr>
<tr>
<td>Dist to Motorable Rd</td>
<td>-0.74 (0.341)</td>
<td>-0.72 (0.342)</td>
<td>-0.74 (0.342)</td>
<td>-0.724 (0.342)</td>
</tr>
<tr>
<td>Dist to Tarmac Rd (x100)</td>
<td>-0.26 (0.069)</td>
<td>-0.27 (0.070)</td>
<td>-0.26 (0.069)</td>
<td>-0.293 (0.071)</td>
</tr>
<tr>
<td>Dist to Extension (x100)</td>
<td>-0.36 (0.116)</td>
<td>-0.34 (0.117)</td>
<td>-0.36 (0.116)</td>
<td>-0.341 (0.117)</td>
</tr>
<tr>
<td>Credit: Tried (x10)</td>
<td>-</td>
<td>0.14 (0.104)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Credit: Unsuccessful Try (x10)</td>
<td>-</td>
<td>-</td>
<td>0.07 (0.199)</td>
<td>-</td>
</tr>
<tr>
<td>Credit: Received (x10)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.18 (0.109)</td>
</tr>
<tr>
<td>P value, Education</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>
Robustness Checks

i. Heckit Selection Equations

ii. ATE, TT, MTE (selection corrected)  
   Excluded Regressor:  
   Distance to closest fertilizer distributor

iii. IV/LATE Estimates

LATE/IV Estimates

Exclusion restriction:  
Use distance to closest fertilizer seller (km) as the excluded IV  
Used for LATE, Heckit and selection corrected ATE, TT and MTE

<table>
<thead>
<tr>
<th>First Stage: Effect of Distance on Probability of Using Hybrid (x100)</th>
<th>Second Stage: Effect of Predicted Hybrid on Yields</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.525 (0.097)</td>
<td>1.452 (0.419)</td>
</tr>
</tbody>
</table>
Marginal Treatment Effects

<table>
<thead>
<tr>
<th>Year</th>
<th>λ Hybrid</th>
<th>λ Non-Hybrid</th>
<th>ATE</th>
<th>TT</th>
<th>MTE Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>-0.833 (0.170)</td>
<td>1.034 (0.724)</td>
<td>2.174</td>
<td>1.138</td>
<td>-1.867 (0.743)</td>
</tr>
<tr>
<td>2000</td>
<td>-0.847 (0.118)</td>
<td>-0.615 (0.353)</td>
<td>0.768</td>
<td>0.898</td>
<td>-0.232 (0.372)</td>
</tr>
<tr>
<td>2004</td>
<td>-0.890 (0.183)</td>
<td>-0.005 (0.156)</td>
<td>1.331</td>
<td>1.014</td>
<td>-0.885 (0.240)</td>
</tr>
</tbody>
</table>

Conclusion: Policy Questions

We saw the policy implications of the other stories in the literature.

Here what are the policy implications? Very different!

You can target policy to be extremely cost effective.

If households have zero returns, policy should not encourage adoption of existing varieties via say extension services.

Farmers that have low returns and use hybrid are not constrained & would benefit from development of new strains.
Role of Infrastructure

Hybrids cause a “reversal of fortune” in that they help those who got the lowest yields from traditional ones.

One way to improve yields (and hence food security), is to develop infrastructure to the areas where hybrid yields are the highest.

Includes encouraging seed and fertilizer suppliers to locate in those areas.

Thank you!