Signals, Similarity and Seeds: Social Learning in the Presence of Imperfect Information and Heterogeneity

Emilia Tjernström

University of California, Davis

November 3, 2014
1 Introduction
   • Motivation
   • Context

2 Research design
   • Data sources
     • RCT
       • Network info
   • Variable definitions
   • Econometrics

3 Results
   • Data
   • Social network results
   • Heterogeneity
Learning & technology adoption

- Greater use of improved technologies could raise productivity and welfare in developing countries
- Returns are typically unknown and stochastic
- Understanding how individuals learn & decide what technologies to use crucial to boosting prosperity
Learning & technology adoption in agriculture

- Agricultural technologies provide a favorable and important context for the study of learning.
- Farmers make production choices in an environment characterized by imperfections, where learning is difficult:
  - financial imperfections: credit constraints and imperfect insurance markets
  - incomplete information about the availability and profitability of new technologies
  - complex and heterogeneous information environment
- Social learning plays a role in diffusion and adoption (Foster & Rosenzweig, 1995; Bandiera & Rasul, 2006; Conley & Udry, 2010; Magnan et al., 2013; Cai et al., 2014; Carter et al., 2014; Adhvaryu, 2014)
Agricultural productivity in SSA: low and stagnant

Figure: Cereal yields in SSA & other regions
Hybrids in Kenya

- Hybrid use is higher than many other SSA countries (40-70%)
- Stagnating maize production partly due to slow replacement of old hybrids
  - 2/3 of farmers grow a hybrid developed in 1986, suited for the Kenyan highlands (Tegemeo, 2010)
  - relevant decision is type of hybrid & this choice is complex
    - many seeds to choose from
    - soil quality varies widely
Farmers face substantial and growing complexity

Figure: Number of maize varieties released in Kenya, 1964 - 2014 and their reported yield capacity
Region exhibits significant heterogeneity in soil quality

Figure: Box plot of Cation Exchange Capacity across sample villages
What I do & summary of results

- Experimental variation in information available to farmers about new tech
  - construct a measure of the signal in individuals’ networks
  - examine how social networks affect familiarity, WTP and adoption of new tech

- Networks matter: they affect
  - familiarity
  - WTP
  - adoption

- Unobserved heterogeneity makes individuals less likely to respond to their peers’ experiences
What I do & summary of results

- Experimental variation in information available to farmers about new tech
  - construct a measure of the signal in individuals’ networks
  - examine how social networks affect familiarity, WTP and adoption of new tech

- Networks matter: they affect
  - familiarity
  - WTP
  - adoption

- Unobserved heterogeneity makes individuals less likely to respond to their peers’ experiences
Impact evaluation

- Large-scale RCT: “Evaluating the socio-economic impacts of Western Seed’s hybrid maize program”
- Western Seed Company (WSC)
  - high-yielding maize hybrids
  - adapted to mid- & low- altitude areas
- Until recently, limited by capacity-constraints
Impact evaluation

- Study villages are in WSC expansion areas
  - no/little information or marketing
  - no/little access to the seeds
  - may have experience with other hybrids
- Cluster-randomized roll-out
  - information about WSC
  - 250g samples of the seeds
    - could plant small experimental plot
    - \(\frac{1}{30}\) of average farmers land
Impact evaluation

- Villages divided into treatment and control clusters
- Sampled farmers in treatment villages received info & samples
- *Main goal:* induce different adoption levels between treatment and control villages
- *Experiment-within-experiment:* variation within treatment villages in the level of experience with the new technology
  - orthogonal to farmer attributes & social network characteristics
## Farmer types

<table>
<thead>
<tr>
<th>Farmer type</th>
<th>Village</th>
<th>Info + sample</th>
<th>Baseline sample</th>
<th>Soil sample</th>
<th>Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directly treated</td>
<td>Treatment</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Indirectly treated</td>
<td>Treatment</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Control</td>
<td>Control</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>
## Farmer types

<table>
<thead>
<tr>
<th>Farmer type</th>
<th>Village</th>
<th>Info + sample</th>
<th>Baseline sample</th>
<th>Soil sample</th>
<th>Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directly treated</td>
<td>Treatment</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Indirectly treated</td>
<td>Treatment</td>
<td>Yes</td>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Control</td>
<td>Control</td>
<td>Yes</td>
<td></td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>
Impact evaluation - timeline

2013

Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

Sampling of villages + hhs
Info sessions & seed samples
Baseline survey

Main season

2014

Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

Phone survey I
Phone survey II
Phone survey III

Network module & learning experiments

**Figure**: RCT timeline

**Tjernström**

**Signals, Similarity and Seeds**
Impact evaluation - timeline

Sampling of villages + hh's
Info sessions & seed samples

Baseline survey

Main season

2013
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

2014
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

Phone survey I

Phone survey II

Network module & learning experiments

Phone survey III

Figure: RCT timeline
Impact evaluation - timeline

Sampling of villages + hhs
Info sessions & seed samples

Baseline survey

Main season

2013
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug

2014
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug

Phone survey I

Network module & learning experiments

Phone survey II

Phone survey III

Figure: RCT timeline
Impact evaluation - timeline

Sampling of villages + hhs
Info sessions & seed samples

2013
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

Main season

2014
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

Phone survey I
Phone survey II
Network module & learning experiments
Phone survey III

Baseline survey

Figure: RCT timeline
Impact evaluation - timeline

Sampling of villages + hhs
Info sessions & seed samples

2013
Jan  Feb  Mar  Apr  May  Jun  Jul  Aug

Main season

Phone survey I

2014
Jan  Feb  Mar  Apr  May  Jun  Jul  Aug

Network module & learning experiments

Phone survey II

Baseline survey

Sep  Oct  Nov  Dec

Short season

Phone survey III

Figure: RCT timeline
Network information

- Additional survey in 20 treatment villages
  - all directly treated hhs
  - random sample of indirectly treated
- 600 farmers invited; 575 (96%) showed up & participated
- Indirectly treated answered additional survey since not in baseline
Different network types

- Information neighbors
- Talk to (about anything, about ag + at different frequencies)
- Economic (microfinance, women’s group, farming group)
- Geographic (walk/bike by, live closest to)
- Information (advice, what seeds they planted/prefer, most similar to you, recommend WSC hybrids)
Tablet network module

Which of these HOUSEHOLDS do you discuss agriculture with?

None / Don’t know
Which of these HOUSEHOLDS do you discuss agriculture with?

None / Don't know
Tablet network module
Tablet network module

Which hhs do you talk to (about ag) MORE than once a YEAR?
Tablet network module

Which hhs do you talk to (about ag) MORE than once a YEAR?
Tablet network module

Which hhs do you talk to (about ag) MORE than once a MONTH?
Tablet network module
Tablet network module
Network definition

- For present analysis, individual $j$ is in person $i$’s social network if person $i$ listed them in *any* of the network questions.

- Many options for defining information networks:
  - *reciprocal*: $i$ mentions $j$ and $j$ mentions $i$
  - *corrected*: remove those who spoke about maize for the first time after treatment.
Network definition

- For present analysis, individual $j$ is in person $i$’s social network if person $i$ listed them in any of the network questions.
- Many options for defining information networks:
  - *reciprocal*: $i$ mentions $j$ and $j$ mentions $i$
  - *corrected*: remove those who spoke about maize for the first time after treatment.
Several recent papers use experimental variation in networks (Carter et al., 2014; Cai et al., 2014; Magnan et al., 2013; Oster & Thornton, 2012)

Unlike earlier observational studies that used innovative measures of information, the experimental studies rely on number of treated in network

- gets around reflection problem (Manski, 1993)
- implicitly assumes 'social influence’ model, rather than social learning
Information signal

- Phone survey with treated - elicit their experience with the technology

1. Actual experience ($y_i$): “How much did you harvest from the sample pack seeds?”

2. Subjective counterfactual ($\tilde{y}_i$): “How much would you have harvested (same weather, input use, etc) if you had planted the seeds you normally grow instead of WSC hybrids?”

- Denote the perceived experimental gains by $\Delta_i$

\[ \Delta_i = \frac{y_i - \tilde{y}_i}{\tilde{y}_i} \]
Information signal

Figure: Distribution of treated farmers’ evaluation of the performance of the hybrid seed samples
Information signal

- The experiences of the farmers in person $i$’s network combine to form a distribution of signals from which she can learn
  - compute the mean and variance of the signals in a respondent’s network

$$
\mu_i = \sum_{j \in N_i} \frac{\Delta_j}{N_i}
$$

$$
\sigma_i = \sum_{j \in N_i} \frac{(\Delta_j - \mu_i)^2}{N_i}
$$

- A higher $\mu_i$ should increase likelihood that farmer $i$ adopts
- A higher $\sigma_i$, i.e. a noisier signal, should decrease farmer $i$’s response to the signal
Information signal

Figure: Distribution of $\mu_i$
Outcome variables

- Familiarity with WSC hybrids
- WTP for WSC hybrids
- Planted a WSC variety
- Planted a non-WSC variety
Familiarity with WSC hybrids

- Indicator variable equal to 1 if respondent is familiar with the technology
- 1\textsuperscript{st} stage of WTP module:
  - respondents shown cards with names of ca. 20 seed varieties
  - asked whether they feel they know enough about the varieties to decide whether or not they would like to plant them
- Measures whether respondent has enough knowledge about WSC hybrid to compare the tech to other seeds?
- Intuitively, have to be familiar with the seed before adopting
  - more restrictive than 'have you heard of WSC hybrids?'
Price-premium based WTP

- 2\textsuperscript{nd} stage of WTP module:
  - rank the seeds with which familiar

- 3\textsuperscript{rd} stage:
  - if ranked a WSC variety over another hybrid, elicited premium
  - add premium to the price of the other hybrid

- Could pick up learning if adoption impacts are limited by liquidity constraints and/or other market imperfections

- Not everyone answers the WTP module
Actual planting behavior

- Planted a WSC variety (0/1)
  - more stringent measure of adoption than other experimental network papers
    - Bandiera & Rasul, 2006; Cai et al, 2014; Oster & Thornton, 2012; Miguel & Kremer, 2004

- Planted a non-WSC hybrid
  - could be 0, positive or negative depending on previous hybrid use and/or spillovers
General specification

\( y_{iv} = f(N_{iv}) + \gamma X_i + \epsilon_{iv} \)

- \( y_{iv} \) is outcome for household \( i \) in village \( v \)
- \( X_i \) is vector of baseline control variables
- \( f(N_{iv}) \) function of information in individual \( i \)'s network
- s.e.'s clustered at village level
General specification

\[ y_{iv} = f(N_{iv}) + \gamma X_i + \varepsilon_{iv} \]

- \( N_{iv} \) represents either
  1. number of treated farmers in farmer \( i \)'s network
  2. first two moments of distribution of experiences reported by treated individuals in her network
- Recent experimental studies typically only consider 1)
These “social influence” models include the number of treated in network in different forms:

- # of treated (Babcock & Hartman, 2010; Oster & Thornton, 2012)
- share of treated (Cai et al., 2014)
- indicator vars for having 1, 2, 3… treated members (Carter et al., 2014)
- dummy for having any treated network members (Magnan et al., 2013)

I use dummies for 1 and “2 or more” treated network members.
Social networks model

- 'Social influence' model:

\[ y_{iv} = \alpha_1 + \beta_k \sum_{k=1}^{K} l_{iv}^k + \gamma_1 x_i + \varepsilon_{iv} \]

where \( K \) in our preferred model is 2+

- Information signal model:

\[ y_{iv} = \alpha_2 + \lambda_k \sum_{k=1}^{2} m_{iv}^k + \gamma_2 x_i + \nu_{iv} \]

\( m_i^k \) denotes the \( k^{th} \) moment of the distribution of signals in person \( i \)'s network
Social networks model

- Estimate most models using OLS
- When outcome variable is WTP for technology, use Tobit as it might be censored at 0
Social networks model

Controls include

- proxies for prior experience with improved tech:
  - dummy for being in a village where the majority of treated do not know where to purchase
  - dummy for having used hybrids & fertilizer

- household characteristics:
  - size of main maize field
  - risk attitudes
  - understanding score from experiments
  - PPI score
  - microfinance participation

- network controls:
  - total network size; signal-regressions also dummies for number of treated links
Heterogeneity

- Cation Exchange Capacity (CEC): summary statistic of soil quality
  - often used to gauge soil fertility
  - varies in sample villages & the extent of variation also varies between villages

- Compute the coefficient of variation (CV) of CEC: measure of unobserved heterogeneity

- Interact $\text{CV}_{CEC}$ with social network variables
### Summary statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>mean</th>
<th>sd</th>
<th>min</th>
<th>max</th>
<th>mean(T) - mean(I)</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Household characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kiswahili spoken at home</td>
<td>0.03</td>
<td>0.18</td>
<td>0</td>
<td>1</td>
<td>-0.001</td>
<td>(-0.06)</td>
</tr>
<tr>
<td>Luhya spoken at home</td>
<td>0.19</td>
<td>0.39</td>
<td>0</td>
<td>1</td>
<td>0.045</td>
<td>(1.41)</td>
</tr>
<tr>
<td>Luo spoken at home</td>
<td>0.78</td>
<td>0.42</td>
<td>0</td>
<td>1</td>
<td>-0.045</td>
<td>(-1.29)</td>
</tr>
<tr>
<td>In womens’ or farm group</td>
<td>0.48</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
<td>0.076*</td>
<td>(1.83)</td>
</tr>
<tr>
<td>In microfinance group</td>
<td>0.25</td>
<td>0.43</td>
<td>0</td>
<td>1</td>
<td>0.009</td>
<td>(0.25)</td>
</tr>
<tr>
<td>General risk taking attitude (0-10)</td>
<td>8.15</td>
<td>2.04</td>
<td>0</td>
<td>10</td>
<td>0.081</td>
<td>(0.47)</td>
</tr>
<tr>
<td>Understanding score, exp. games</td>
<td>0.74</td>
<td>0.34</td>
<td>0</td>
<td>1</td>
<td>-0.024</td>
<td>(-0.85)</td>
</tr>
<tr>
<td>PPI score (0-100)</td>
<td>44.49</td>
<td>12.41</td>
<td>14</td>
<td>84</td>
<td>1.409</td>
<td>(1.35)</td>
</tr>
</tbody>
</table>

* t statistics in parentheses, standard errors clustered at the village level
* * p<.1, ** p<.05, *** p<.01

**Table**: Summary statistics
## Summary statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>mean</th>
<th>sd</th>
<th>min</th>
<th>max</th>
<th>mean(T) - mean(I)</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Agricultural characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size of main maize field (acres)</td>
<td>1.30</td>
<td>1.16</td>
<td>.07</td>
<td>10</td>
<td>0.201**</td>
<td>(2.16)</td>
</tr>
<tr>
<td>Nr. of seasons used fertilizer, 4 years</td>
<td>2.57</td>
<td>3.33</td>
<td>0</td>
<td>8</td>
<td>0.479*</td>
<td>(1.71)</td>
</tr>
<tr>
<td>Nr. of seasons used hybrids, 4 years</td>
<td>3.32</td>
<td>3.33</td>
<td>0</td>
<td>8</td>
<td>-0.059</td>
<td>(-0.21)</td>
</tr>
<tr>
<td><strong>Network characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nr. of relatives</td>
<td>2.43</td>
<td>2.23</td>
<td>0</td>
<td>12</td>
<td>0.070</td>
<td>(0.38)</td>
</tr>
<tr>
<td>Nr. of treated relatives</td>
<td>1.31</td>
<td>1.39</td>
<td>0</td>
<td>8</td>
<td>0.080</td>
<td>(0.69)</td>
</tr>
<tr>
<td>Nr. of links (all)</td>
<td>7.05</td>
<td>3.92</td>
<td>0</td>
<td>29</td>
<td>0.344</td>
<td>(1.08)</td>
</tr>
<tr>
<td>Nr. of treated links (all)</td>
<td>4.08</td>
<td>2.51</td>
<td>0</td>
<td>20</td>
<td>0.549***</td>
<td>(2.69)</td>
</tr>
<tr>
<td>Nr. of reciprocal links (all)</td>
<td>3.29</td>
<td>2.50</td>
<td>0</td>
<td>22</td>
<td>0.409**</td>
<td>(2.01)</td>
</tr>
<tr>
<td>Nr. of treated reciprocal links (all)</td>
<td>1.93</td>
<td>1.71</td>
<td>0</td>
<td>15</td>
<td>0.435***</td>
<td>(3.15)</td>
</tr>
<tr>
<td>Nr. of links in corrected network</td>
<td>6.73</td>
<td>3.78</td>
<td>0</td>
<td>29</td>
<td>0.154</td>
<td>(0.50)</td>
</tr>
<tr>
<td>Nr. of treated links, corrected network</td>
<td>3.85</td>
<td>2.41</td>
<td>0</td>
<td>19</td>
<td>0.400**</td>
<td>(2.03)</td>
</tr>
</tbody>
</table>

$t$ statistics in parentheses, standard errors clustered at the village level
* p<.1, ** p<.05, *** p<.01
Balance on observables

- Require that treatment induced exogenous variation in number of treated network members in a given individual’s network
  - conditional on individual $i$’s total number of links (total network size), the number of treated links was randomized
  - test the validity this assumption by regressing baseline characteristics on number of treated links (controlling for total network size)
- Do this separately for treated & indirectly treated
- Test using 3 different network definitions
## Balance on observables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Relatives</th>
<th>Corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>I</td>
</tr>
<tr>
<td><strong>Household characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In womens' or farm group</td>
<td>-0.009</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(-0.20)</td>
<td>(0.23)</td>
</tr>
<tr>
<td>In microfinance group</td>
<td>-0.047*</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(-1.90)</td>
<td>(-0.07)</td>
</tr>
<tr>
<td>General risk taking perception (0-10)</td>
<td>-0.089</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>(-0.50)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>Understanding score, exp. games</td>
<td>-0.010</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td>(-0.42)</td>
<td>(1.33)</td>
</tr>
<tr>
<td>Sum of core 10 PPI scores (0-100)</td>
<td>-0.506</td>
<td>1.248</td>
</tr>
<tr>
<td></td>
<td>(-0.68)</td>
<td>(1.09)</td>
</tr>
</tbody>
</table>

*t statistics in parentheses, standard errors clustered at the village level

* p<.1, ** p<.05, *** p<.01

**Table**: Regression of baseline vars on nr. of treated links
### Balance on observables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Relatives</th>
<th>Corrected</th>
<th>Relatives</th>
<th>Corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T</td>
<td>I</td>
<td>T</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Agricultural characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size of main maize field (acres)</td>
<td>-0.026 (-0.27)</td>
<td>0.024 (0.35)</td>
<td>-0.029 (-0.55)</td>
<td>-0.038 (-0.69)</td>
</tr>
<tr>
<td>Nr. of seasons used fertilizer, 4 years</td>
<td>0.440 (1.37)</td>
<td>0.271 (1.07)</td>
<td>0.303 (1.56)</td>
<td>0.536*** (3.21)</td>
</tr>
<tr>
<td>Nr. of seasons used hybrids, 4 years</td>
<td>0.334 (1.26)</td>
<td>0.882*** (2.92)</td>
<td>0.244 (1.32)</td>
<td>0.628*** (3.88)</td>
</tr>
</tbody>
</table>

_t statistics in parentheses, standard errors clustered at the village level_

* _p<.1, ** _p<.05, *** _p<.01

**Table**: Regression of baseline vars on nr. of treated links
### Familiarity, social influence model

*Dep. variable: Familiar with WSC hybrid?*

<table>
<thead>
<tr>
<th>Nr. of treated links</th>
<th>Panel A - Treated</th>
<th></th>
<th>Panel A - Indirectly treated</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 treated link</td>
<td>0.20 (0.2)</td>
<td>0.097 (0.3)</td>
<td>0.29 (0.3)</td>
<td>0.020 (0.1)</td>
</tr>
<tr>
<td>2+ treated links</td>
<td>0.31 (0.2)</td>
<td>0.50* (0.3)</td>
<td>0.47* (0.3)</td>
<td>0.082 (0.2)</td>
</tr>
<tr>
<td>Network size</td>
<td>0.0071 (0.006)</td>
<td>0.13 (0.1)</td>
<td>0.0042 (0.007)</td>
<td>0.013 (0.01)</td>
</tr>
<tr>
<td>(1 treated)*(nw. size)</td>
<td>-0.036 (0.1)</td>
<td></td>
<td>-0.23*** (0.07)</td>
<td></td>
</tr>
<tr>
<td>(2+ treated)*(nw. size)</td>
<td>-0.12 (0.1)</td>
<td></td>
<td>-0.18** (0.06)</td>
<td></td>
</tr>
<tr>
<td>On-farm trial outcome</td>
<td>0.00067 (0.03)</td>
<td></td>
<td></td>
<td>0.00016 (0.002)</td>
</tr>
<tr>
<td>(On-farm trial outcome)^2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Additional covars</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>319</td>
<td>319</td>
<td>217</td>
<td>255</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.078</td>
<td>0.083</td>
<td>0.087</td>
<td>0.229</td>
</tr>
</tbody>
</table>

In both panels: standard errors in parentheses; s.e.'s clustered at the village level; * p<.1, ** p<.05, *** p<.01

*Network definition used:* individual $j$ is in person $i$’s network if person $i$ listed them in *any* of the network questions.

**Table:** Social network effects on farmer familiarity with WSC hybrids

---

Tjernström | Signals, Similarity and Seeds
### Familiarity, information signal model

*(Dep. variable: Familiar with WSC hybrid?)*

<table>
<thead>
<tr>
<th>Panel B - Signal in nw</th>
<th>Treated 1</th>
<th>Treated 2</th>
<th>Indirectly treated 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. signal in nw.</td>
<td>0.022</td>
<td>-0.027</td>
<td>0.00024 (0.01)</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.04)</td>
<td></td>
</tr>
<tr>
<td>Variance of signal in nw.</td>
<td>-0.0000016</td>
<td>0.0022</td>
<td>-0.0046*** (0.0010)</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>Network size</td>
<td>0.0066</td>
<td>0.0019</td>
<td>0.014 (0.01)</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.007)</td>
<td></td>
</tr>
<tr>
<td>On-farm trial outcome</td>
<td>0.0073</td>
<td></td>
<td>0.0073 (0.03)</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(On$-farm trial outcome$)^2$</td>
<td></td>
<td>-0.00017</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Additional covars</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>294</td>
<td>202</td>
<td>227</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.042</td>
<td>0.006</td>
<td>0.238</td>
</tr>
</tbody>
</table>

In both panels: standard errors in parentheses; s.e.’s clustered at the village level; * p<.1, ** p<.05, *** p<.01

Network definition used: individual $j$ is in person $i$’s network if person $i$ listed them in any of the network questions.

**Table**: Social network effects on farmer familiarity with WSC hybrids
## WTP, social influence model

*(Dep. variable: Willingness to pay for WSC hybrid)*

<table>
<thead>
<tr>
<th>Panel A - treated links</th>
<th>Nr. of treated links</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 treated link</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 treated link</td>
<td>83.0</td>
<td>84.1</td>
<td>314.9***</td>
</tr>
<tr>
<td></td>
<td>(77.0)</td>
<td>(126.7)</td>
<td>(73.9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2+ treated links</td>
<td>116.8**</td>
<td>96.1</td>
<td>263.0***</td>
</tr>
<tr>
<td></td>
<td>(51.7)</td>
<td>(108.9)</td>
<td>(66.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Network size</td>
<td>2.40</td>
<td>4.13</td>
<td>9.49</td>
</tr>
<tr>
<td></td>
<td>(3.8)</td>
<td>(4.5)</td>
<td>(9.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>On-farm trial outcome</td>
<td>26.6</td>
<td>(18.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(On-farm trial outcome)^2</td>
<td>-1.80</td>
<td>(1.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Additional covars</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Observations</td>
<td>224</td>
<td>173</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>Adjusted R^2</td>
<td>0.064</td>
<td>0.087</td>
<td>0.075</td>
</tr>
</tbody>
</table>

In both panels: standard errors in parentheses; s.e.'s clustered at the village level; * p<.1, ** p<.05, *** p<.01

*Network definition used:* individual j is in person i’s network if person i listed them in *any* of the network questions.

| Table : Social network effects on farmer WTP for WSC hybrids | Tjernström | Signals, Similarity and Seeds |
### WTP, information signal model

*(Dep. variable: Willingness to pay for WSC hybrid)*

<table>
<thead>
<tr>
<th>Tobit regression</th>
<th>Treated</th>
<th>Indirectly treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel B - Signal in nw</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Avg. signal in nw.</td>
<td>31.0**</td>
<td>25.6</td>
</tr>
<tr>
<td>(14.2)</td>
<td>(16.7)</td>
<td>(19.8)</td>
</tr>
<tr>
<td>Variance of signal in nw.</td>
<td>-1.55**</td>
<td>-1.03</td>
</tr>
<tr>
<td>(0.8)</td>
<td>(0.9)</td>
<td>(6.1)</td>
</tr>
<tr>
<td>Network size</td>
<td>3.92</td>
<td>5.78</td>
</tr>
<tr>
<td>(4.2)</td>
<td>(5.1)</td>
<td>(8.6)</td>
</tr>
<tr>
<td>On-farm trial outcome</td>
<td>30.9</td>
<td>(21.8)</td>
</tr>
<tr>
<td>(On-farm trial outcome)$^2$</td>
<td>-2.04</td>
<td>(1.4)</td>
</tr>
<tr>
<td>Additional covars</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>215</td>
<td>168</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>227.2***</td>
<td>223.4***</td>
</tr>
</tbody>
</table>

In both panels: standard errors in parentheses; s.e.’s clustered at the village level; * p<.1, ** p<.05, *** p<.01

*Network definition used*: individual $j$ is in person $i$’s network if person $i$ listed them in *any* of the network questions.
## WSC hybrid adoption, social influence model

*Dep. variable: Planted WSC hybrid?*

<table>
<thead>
<tr>
<th>Panel A -</th>
<th>Treated</th>
<th>Indirectly treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nr. of treated links</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1 treated link</td>
<td>0.35***</td>
<td>0.32***</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>2+ treated links</td>
<td>0.13**</td>
<td>0.16*</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>Network size</td>
<td>0.0066</td>
<td>0.0051</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>On-farm trial outcome</td>
<td>0.039</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td></td>
</tr>
<tr>
<td>(On-farm trial outcome)$^2$</td>
<td>-0.0029*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Additional covars</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>319</td>
<td>217</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.083</td>
<td>0.073</td>
</tr>
</tbody>
</table>

In both panels: standard errors in parentheses; s.e.'s clustered at the village level; * p<.1, ** p<.05, *** p<.01

Network definition used: individual $j$ is in person $i$'s network if person $i$ listed them in *any* of the network questions.

### Table: Social network effects on probability of planting a WSC hybrid
### WSC hybrid adoption, information signal model

*(Dep. variable: Planted WSC hybrid?)*

<table>
<thead>
<tr>
<th>Panel B - Signal in nw</th>
<th>Treated</th>
<th>Indirectly treated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Avg. signal in nw.</td>
<td>-0.023</td>
<td>-0.032</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Variance of signal in nw.</td>
<td>0.0034</td>
<td>0.0044**</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Network size</td>
<td>0.0065</td>
<td>0.0048</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>On-farm trial outcome</td>
<td>0.042</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td></td>
</tr>
<tr>
<td>(On-farm trial outcome)^2</td>
<td></td>
<td>-0.0029*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.001)</td>
</tr>
<tr>
<td>Additional covars</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>294</td>
<td>202</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.088</td>
<td>0.072</td>
</tr>
</tbody>
</table>

In both panels: standard errors in parentheses; s.e.’s clustered at the village level; * $p<.1$, ** $p<.05$, *** $p<.01$  

*Network definition used:* individual $j$ is in person $i$’s network if person $i$ listed them in *any* of the network questions.
Planted other hybrid, social influence model

( Dep. variable: Planted a non-WSC hybrid?)

<table>
<thead>
<tr>
<th>Panel A -</th>
<th>Treated</th>
<th>Indirectly treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nr. of treated links</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1 treated link</td>
<td>-0.35*</td>
<td>-0.21</td>
</tr>
<tr>
<td></td>
<td>(0.2)</td>
<td>(0.2)</td>
</tr>
<tr>
<td>2+ treated links</td>
<td>-0.19</td>
<td>-0.14</td>
</tr>
<tr>
<td></td>
<td>(0.1)</td>
<td>(0.1)</td>
</tr>
<tr>
<td>Network size</td>
<td>0.0080</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>On-farm trial outcome</td>
<td>0.074**</td>
<td>(0.03)</td>
</tr>
<tr>
<td>(On-farm trial outcome)$^2$</td>
<td>-0.0034</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Additional covars</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>319</td>
<td>217</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.166</td>
<td>0.128</td>
</tr>
</tbody>
</table>

In both panels: standard errors in parentheses; s.e.'s clustered at the village level; * p<.1, ** p<.05, *** p<.01

Network definition used: individual $j$ is in person $i$'s network if person $i$ listed them in any of the network questions.

Table: Social network effects on probability of planting a non-WSC hybrid
Table: Social network effects on probability of planting a non-WSC hybrid

(Dep. variable: Planted a non-WSC hybrid?)

<table>
<thead>
<tr>
<th>Panel B - Signal in nw</th>
<th>Treated</th>
<th>Indirectly treated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Avg. signal in nw.</td>
<td>0.027</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Variance of signal in nw.</td>
<td>-0.0040*</td>
<td>-0.0037*</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Network size</td>
<td>0.0089</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>On-farm trial outcome</td>
<td>0.077**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td></td>
</tr>
<tr>
<td>(On-farm trial outcome)^2</td>
<td>-0.0035</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>Additional covars</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>294</td>
<td>202</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.170</td>
<td>0.110</td>
</tr>
</tbody>
</table>

In both panels: standard errors in parentheses; s.e.’s clustered at the village level; * p<.1, ** p<.05, *** p<.01

Network definition used: individual j is in person i’s network if person i listed them in any of the network questions.
Familiarity

**Figure**: How impact of avg. signal in nw. varies with heterogeneity
Figure: How impact of avg. signal in nw. varies with heterogeneity
WSC adoption

Figure: How impact of avg. signal in nw. varies with heterogeneity
Other hybrid

Figure: How impact of avg. signal in nw. varies with heterogeneity
Conclusion

- Use experimental variation in information available through networks to study what farmers learn from their social networks.
- Farmers talk and learn from each other BUT heterogeneity that is unobserved to farmers makes them rely less on information from their peers.
- Can help us understand why some innovations diffuse slowly.
- Can inform policy:
  - when will broad-based extension programs be successful?
  - when do we need to promote individual learning?
- Also useful for thinking about other stochastic technologies.