# Can Farmers Create Efficient Information Networks?

## Experimental Evidence from Rural India

Stefano Caria
Marcel Fafchamps

June 2015


Diffusion of information in social networks is often incomplete.
(Banerjee et al. 2013, Berg et al. 2015)

The structure of networks can limit diffusion.
(Centola 2011, Golub and Jackson 2012)

$\Rightarrow$ Why do network structures that limit diffusion emerge?

Diffusion of information in social networks is often incomplete.
(Banerjee et al. 2013, Berg et al. 2015)

The structure of networks can limit diffusion.
(Centola 2011, Golub and Jackson 2012)

$\Rightarrow$ Why do network structures that limit diffusion emerge?

Diffusion of information in social networks is often incomplete.
(Banerjee et al. 2013, Berg et al. 2015)

The structure of networks can limit diffusion.
(Centola 2011, Golub and Jackson 2012)

$\Rightarrow$ Why do network structures that limit diffusion emerge?

When links are chosen unilaterally, network structures that guarantee full diffusion are achieved after repeated play.
(Bala and Goyal 2000)

We play this game with a population for which we have evidence of limited information diffusion: farmers in a developing country.

1. Can farmers create efficient networks?
2. Does group categorisation affect network structure and efficiency?

When links are chosen unilaterally, network structures that guarantee full diffusion are achieved after repeated play.
(Bala and Goyal 2000)

We play this game with a population for which we have evidence of limited information diffusion: farmers in a developing country.

1. Can farmers create efficient networks?
2. Does group categorisation affect network structure and efficiency?

When links are chosen unilaterally, network structures that guarantee full diffusion are achieved after repeated play.
(Bala and Goyal 2000)

We play this game with a population for which we have evidence of limited information diffusion: farmers in a developing country.

1. Can farmers create efficient networks?
2. Does group categorisation affect network structure and efficiency?

R1 Farmers achieve much lower network efficiency that what the simple rules would guarantee

R2 Links to the most popular farmer in the network are mostly responsible for the efficiency loss

R3 Information about group membership affects network structure, but not network efficiency

R1 Farmers achieve much lower network efficiency that what the simple rules would guarantee

R2 Links to the most popular farmer in the network are mostly responsible for the efficiency loss

R3 Information about group membership affects network structure, but not network efficiency

R1 Farmers achieve much lower network efficiency that what the simple rules would guarantee

R2 Links to the most popular farmer in the network are mostly responsible for the efficiency loss

R3 Information about group membership affects network structure, but not network efficiency

R1 Farmers achieve much lower network efficiency that what the simple rules would guarantee

R2 Links to the most popular farmer in the network are mostly responsible for the efficiency loss

R3 Information about group membership affects network structure, but not network efficiency

We contribute to the study of information diffusion in networks

- Foster Rosenzweig 1996, Munshi 2004, Bandiera Rasul 2006, Conley Udry 2010, Banerjee et al 2012, Ben Yishay Mobarak 2012, Berg et al 2013

We provide new evidence on how rules of thumb,
and social norms and preferences affect network structure

- Bala Goyal 2000, Akerlof Kranton 2000, Jackson Golub 2009, Berg et al 2013, Kosfeld Falk 2012, Goeree et al 2009

We contribute to the study of information diffusion in networks

- Foster Rosenzweig 1996, Munshi 2004, Bandiera Rasul 2006, Conley Udry 2010, Banerjee et al 2012, Ben Yishay Mobarak 2012, Berg et al 2013

We provide new evidence on how rules of thumb,
and social norms and preferences affect network structure

- Bala Goyal 2000, Akerlof Kranton 2000, Jackson Golub 2009, Berg et al 2013, Kosfeld Falk 2012, Goeree et al 2009

# Outline

1. Design
2. Predictions
3. Results
4. Conclusion

# Players unilaterally create a one-way-flow network

- Subjects are divided into groups of 6
- They form links with others in the group
  - Game is divided in 2 rounds
  - In each round, every player plays once, in random, unknown order
  - Only 1 link allowed. In round 2 players rewire old link
- One player is randomly drawn to win a monetary prize
- Players directly or indirectly connected in the final network to the winner also win the prize

# Players unilaterally create a one-way-flow network

- Subjects are divided into groups of 6
- They form links with others in the group
  - Game is divided in 2 rounds
  - In each round, every player plays once, in random, unknown order
  - Only 1 link allowed. In round 2 players rewire old link
- One player is randomly drawn to win a monetary prize
- Players directly or indirectly connected in the final network to the winner also win the prize

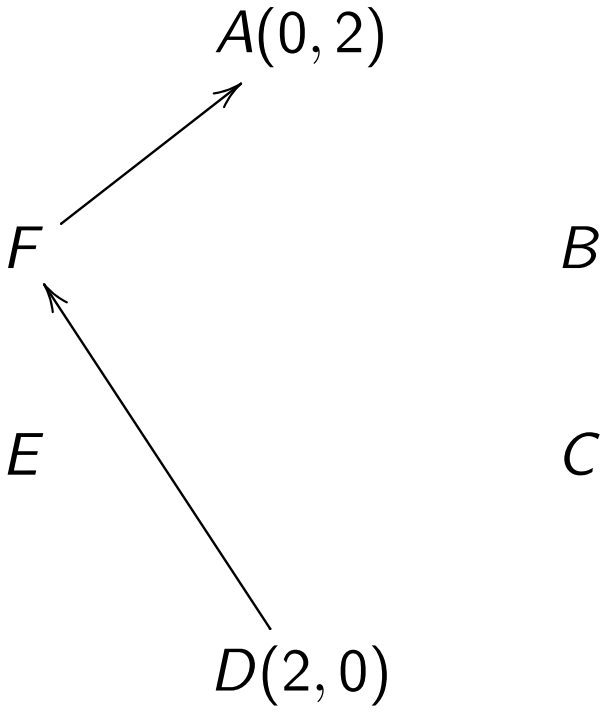
# Players unilaterally create a one-way-flow network

- Subjects are divided into groups of 6
- They form links with others in the group
  - Game is divided in 2 rounds
  - In each round, every player plays once, in random, unknown order
  - Only 1 link allowed. In round 2 players rewire old link
- One player is randomly drawn to win a monetary prize
- Players directly or indirectly connected in the final network to the winner also win the prize

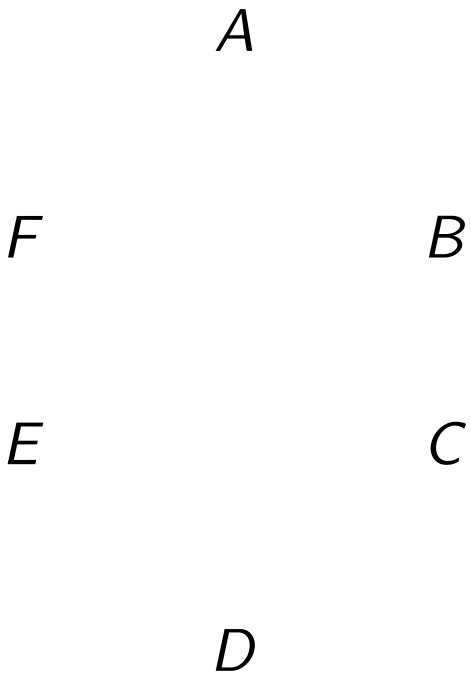
# Players unilaterally create a one-way-flow network

- Subjects are divided into groups of 6
- They form links with others in the group
  - Game is divided in 2 rounds
  - In each round, every player plays once, in random, unknown order
  - Only 1 link allowed. In round 2 players rewire old link
- One player is randomly drawn to win a monetary prize
- Players directly or indirectly connected in the final network to the winner also win the prize

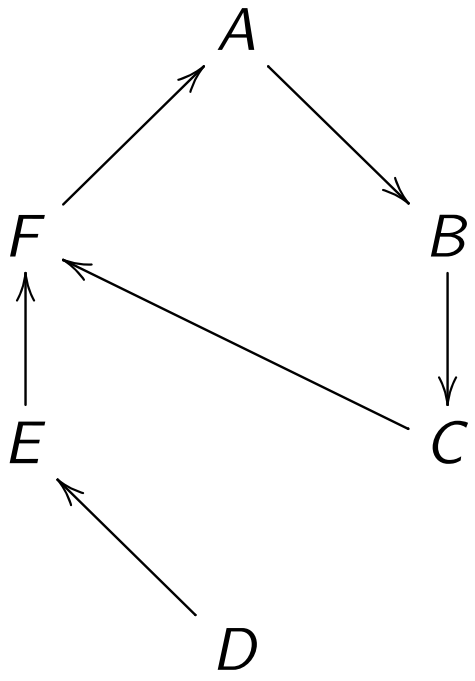
# Reach, in-reach



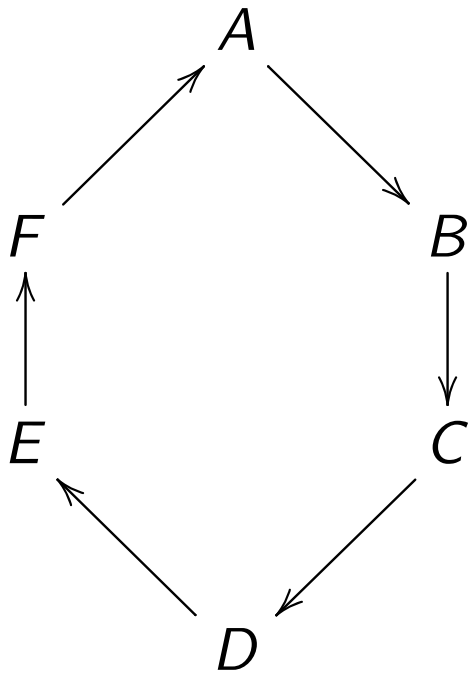
# What will an efficient information network look like?



# An inefficient network



# The efficient cycle network



# A simple measure of efficiency

$$\text{Efficiency}_g = \frac{\frac{1}{n}\sum_{i=1}^{n} \text{reach}_i}{5} \tag{1}$$

# Outline

1 Design

2 Predictions

3 Results

4 Conclusion

# In Treatment 1 players observe the information of others

- Each player i can form a link towards a player j in the group
- Myopic, selfish best response: select the person with the most information, ie. the **maximum reach**
- If every player myopically best responds, the network configuration converges towards the cycle
  - This happens within 2 rounds in (almost) all simulated cases

# In Treatment 1 players observe the information of others

- Each player i can form a link towards a player j in the group
- Myopic, selfish best response: select the person with the most information, ie. the **maximum reach**
- If every player myopically best responds, the network configuration converges towards the cycle
  - This happens within 2 rounds in (almost) all simulated cases

# In Treatment 1 players observe the information of others

- Each player i can form a link towards a player j in the group
- Myopic, selfish best response: select the person with the most information, ie. the **maximum reach**
- If every player myopically best responds, the network configuration converges towards the cycle
  - This happens within 2 rounds in (almost) all simulated cases

# An example

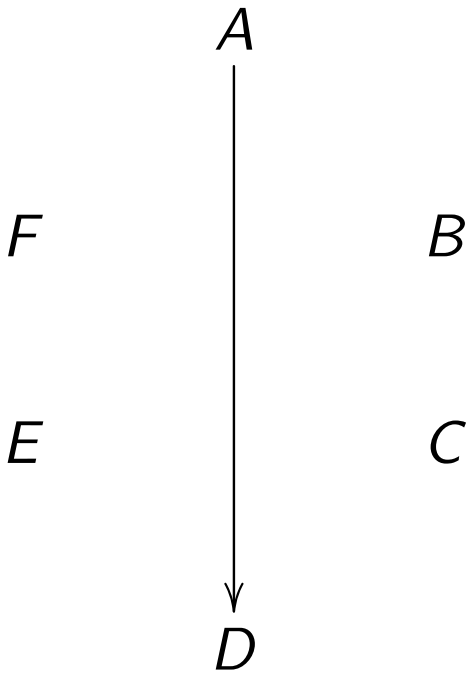



Figure : Move 1

# An example

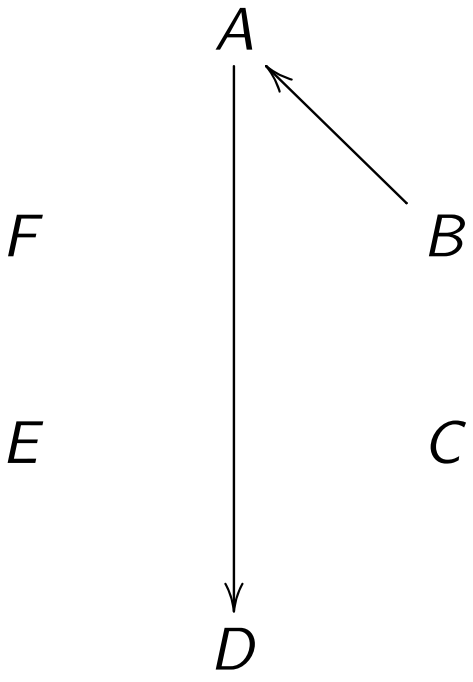



Figure : Move 2

# An example

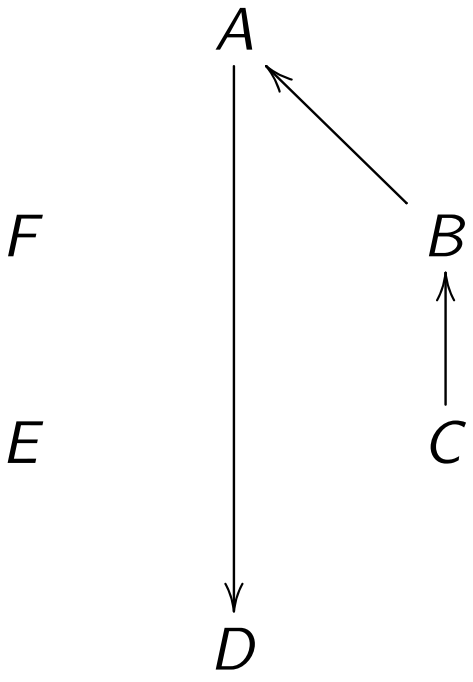



Figure : Move 3

# An example

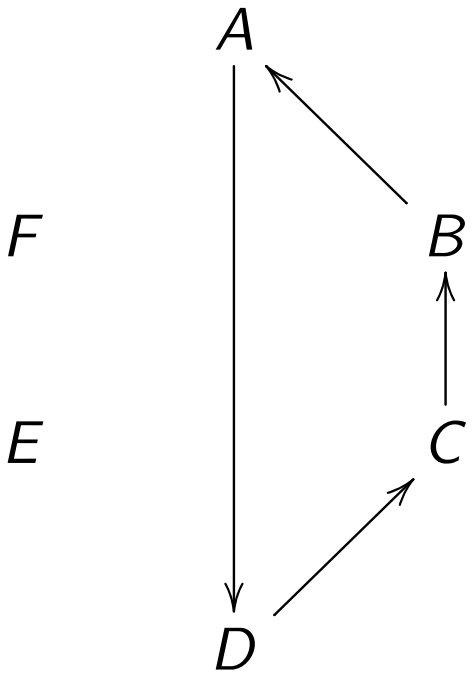



Figure : Move 4

# An example

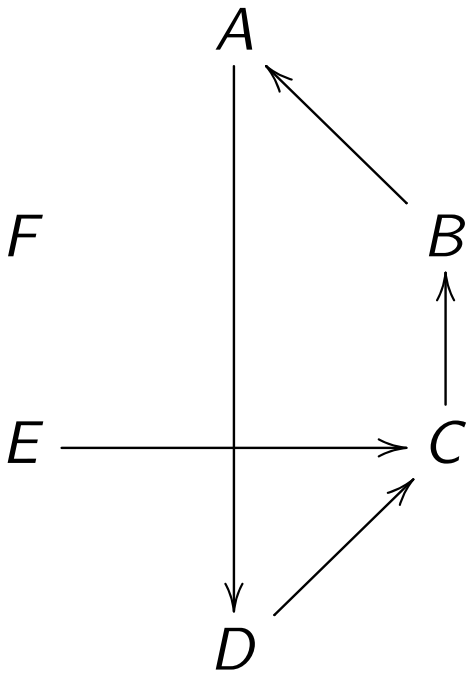



Figure : Move 5

# An example

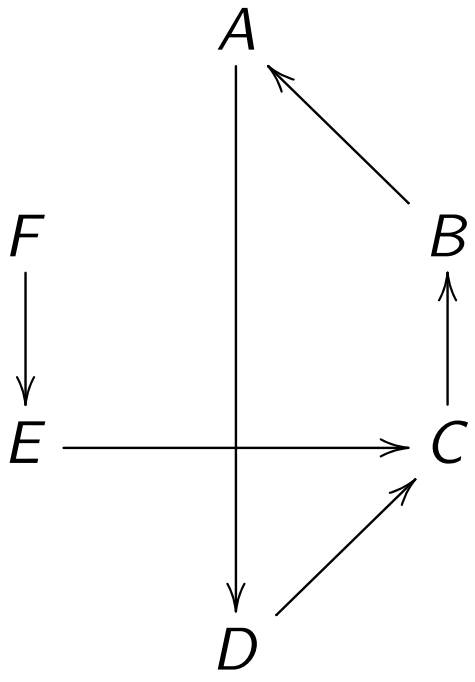



Figure : Move 6

# An example

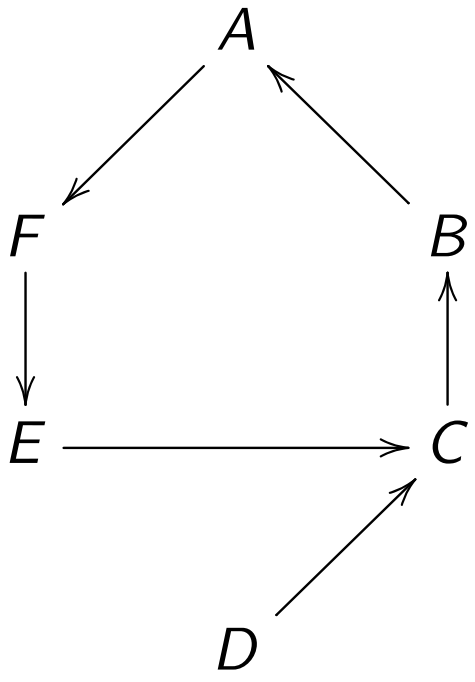



Figure : Move 7

# The cycle!

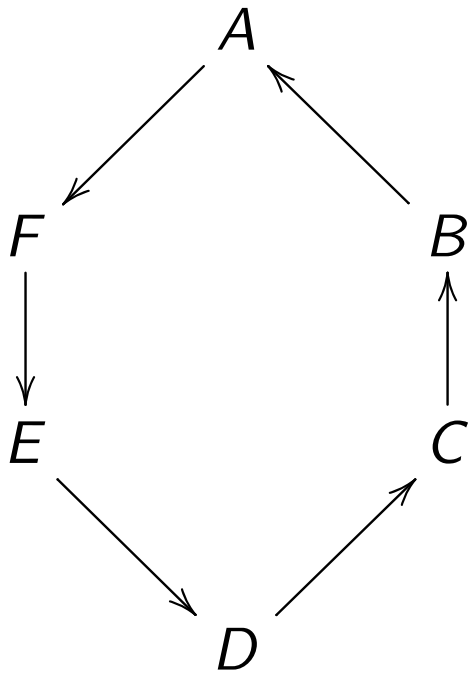



Figure : Move 11

# In T2 players share own information with others

- Each player i can form a link from j to i
- If social welfare maximiser: pick player who passes prize to the largest number of people, ie **maximum in-reach**
  - Network converges towards the cycle in (almost) all cases
- If Rawlsian: **minimum reach**

# In T2 players share own information with others

- Each player i can form a link from j to i
- If social welfare maximiser: pick player who passes prize to the largest number of people, ie **maximum in-reach**
  - Network converges towards the cycle in (almost) all cases
- If Rawlsian: **minimum reach**

# In T2 players share own information with others

- Each player i can form a link from j to i
- If social welfare maximiser: pick player who passes prize to the largest number of people, ie **maximum in-reach**
  - Network converges towards the cycle in (almost) all cases
- If Rawlsian: **minimum reach**

# In T2 players share own information with others

- Each player i can form a link from j to i
- If social welfare maximiser: pick player who passes prize to the largest number of people, ie **maximum in-reach**
  - Network converges towards the cycle in (almost) all cases
- If Rawlsian: **minimum reach**

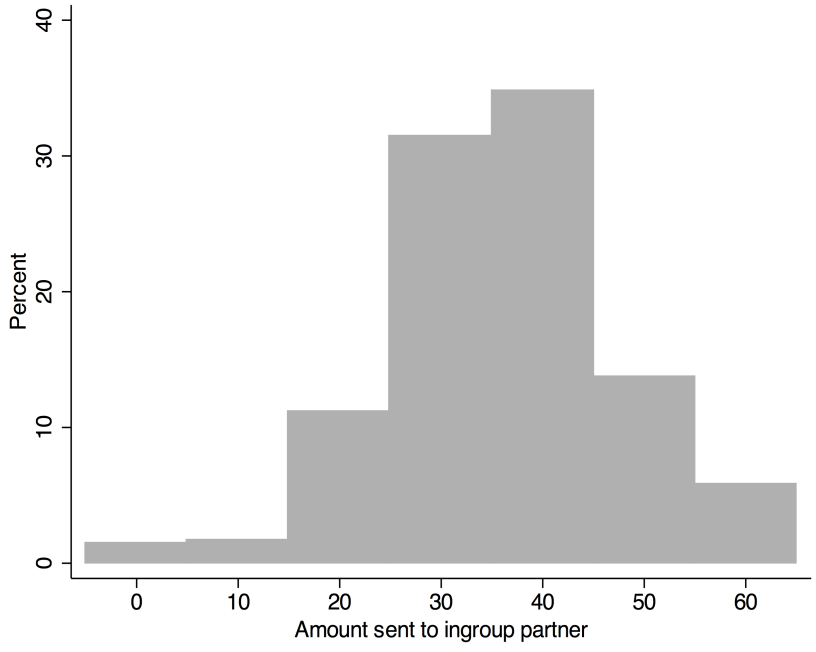
# The effects of group categorisation

- We impose exogenous social identities, create saliency, and make them public in half of the sessions
- Akerlof Kranton (2000) suggest individuals get positive utility from following group prescriptions

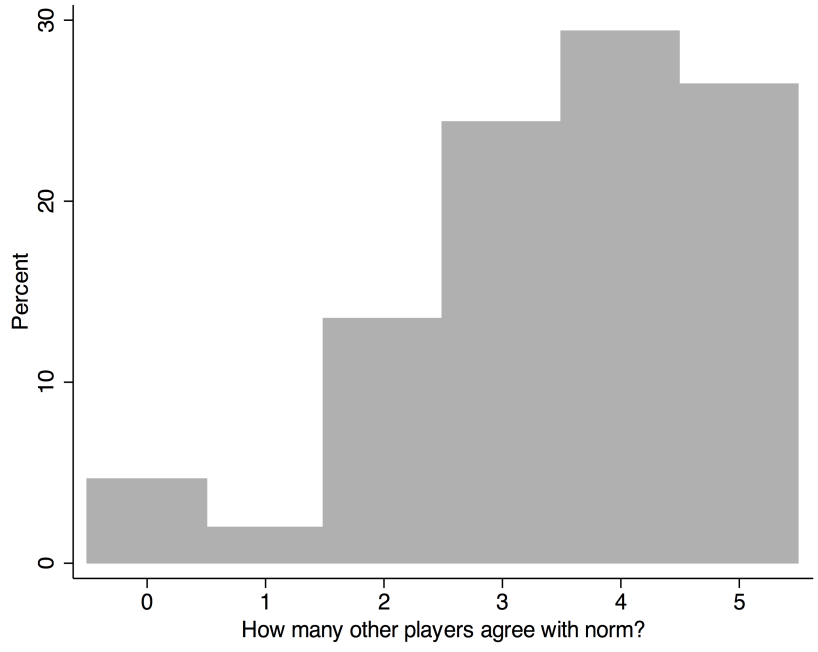
# The effects of group categorisation

- We impose exogenous social identities, create saliency, and make them public in half of the sessions
- Akerlof Kranton (2000) suggest individuals get positive utility from following group prescriptions

# Group salience and in-group norms




(a) Monetary allocation




(b) Normative expectation

# Outline

1 Design

2 Predictions

3 Results

4 Conclusion

# The experiment

- Villages from 1.30 to 3hr from Pune, Maharasthra
  - Random door-to-door sampling of adult male farmers
- 486 participants in 81 sessions
- Balance on the available set of covariates
- 80 percent made at most 2 mistake out of the 7/8 understanding questions

# The experiment

- Villages from 1.30 to 3hr from Pune, Maharasthra
  - Random door-to-door sampling of adult male farmers
- 486 participants in 81 sessions
- Balance on the available set of covariates
- 80 percent made at most 2 mistake out of the 7/8 understanding questions

# The experiment

- Villages from 1.30 to 3hr from Pune, Maharasthra
  - Random door-to-door sampling of adult male farmers
- 486 participants in 81 sessions
- Balance on the available set of covariates
- 80 percent made at most 2 mistake out of the 7/8 understanding questions

- We use non-parametric Wilcoxon rank sum tests for session level outcomes

- and dyadic regression analysis for individual data:

$$\text{link}_{ij,r} = \alpha + \beta \textbf{Network Position}_{j,r} + \gamma \mathbf{D}_{ij} + \delta \text{round}_r + e_{ij,r}$$

- We use non-parametric Wilcoxon rank sum tests for session level outcomes

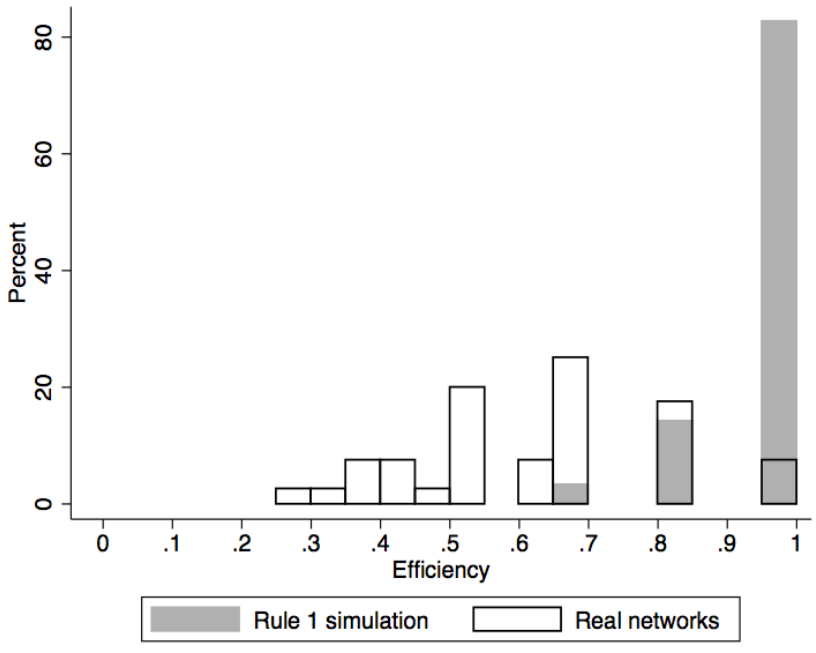
- and dyadic regression analysis for individual data:

$$\text{link}_{ij,r} = \alpha + \beta \textbf{Network Position}_{j,r} + \gamma \mathbf{D}_{ij} + \delta \text{round}_r + e_{ij,r}$$

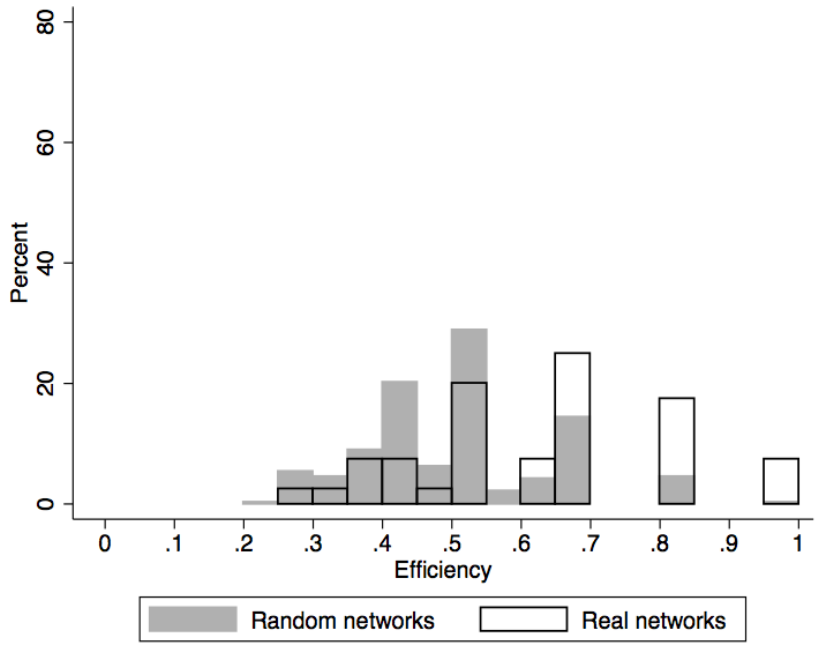
Figure : A session of the experiment

# **Result 1**: Efficiency is 64 percent, significantly lower than what the simple rules would achieve

Figure : Efficiency in no-identity sessions and in simulated networks




(a) Rank sum test: Z= 12.08, p< .001




(b) Rank sum test: Z= 4.62, p< .001

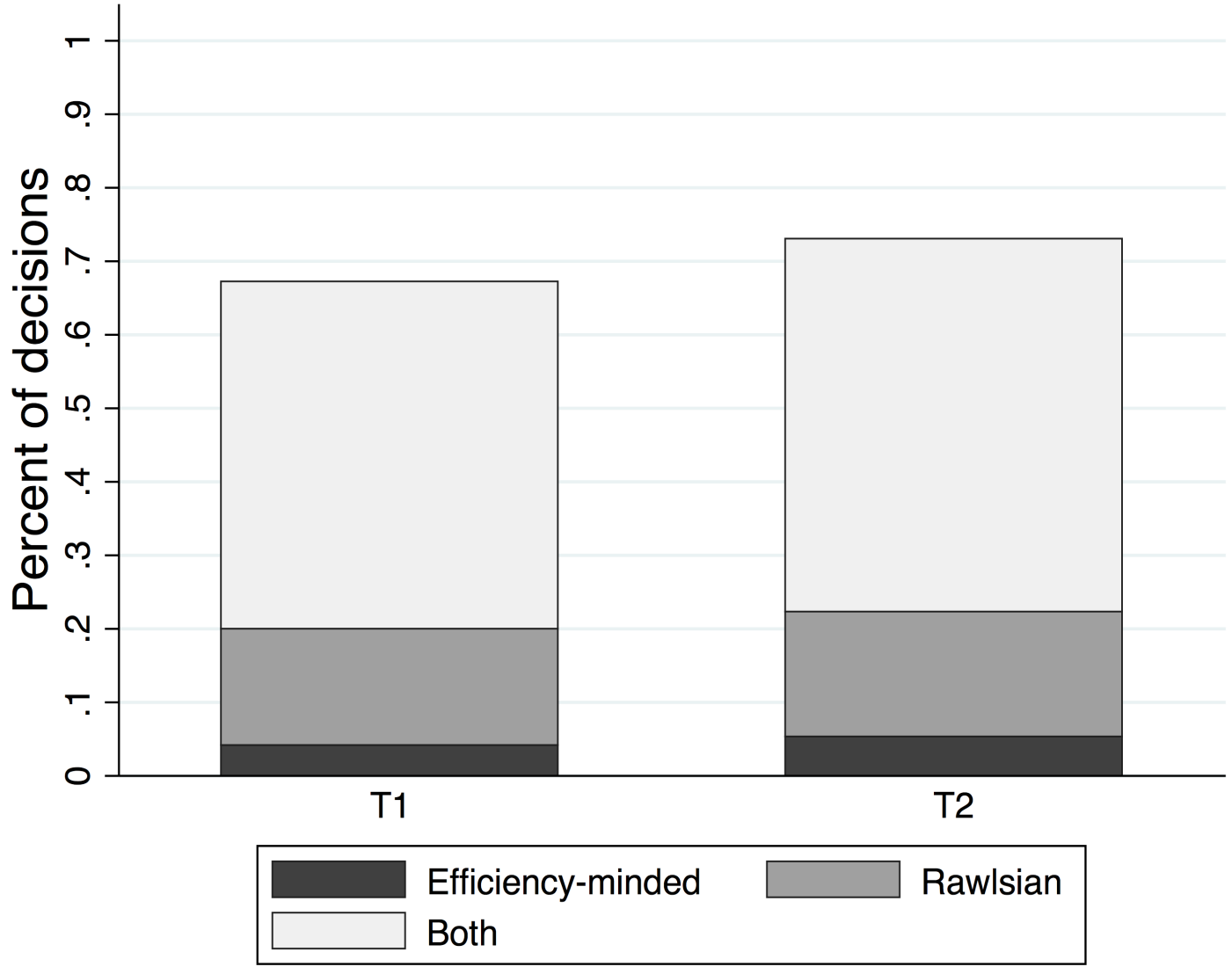
# **Result 2**: Reach and in-reach predict new links

Table : Dyadic linear probability model

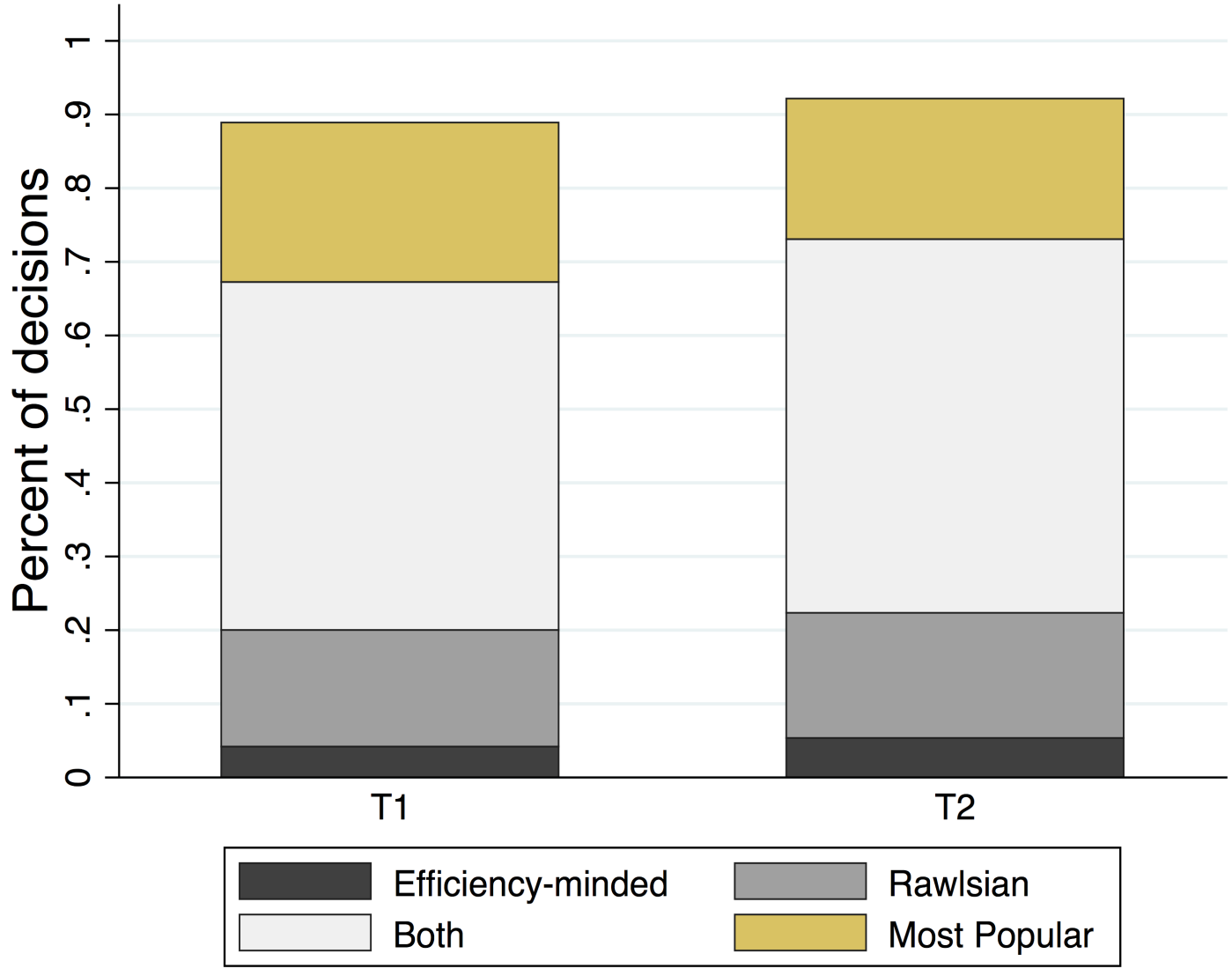
| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **Panel a** | | | | |
| max reach$_j$ | .132 | .130 | | |
| | (.001)*** | (.001)*** | | |
| min in-reach$_j$ | .018 | .016 | | |
| | (.235) | (.314) | | |
| max in-reach$_j$ | | | .111 | .120 |
| | | | (.004)*** | (.002)*** |
| min reach$_j$ | | | .073 | .066 |
| | | | (.04)** | (.072)* |
| Sessions | T1 | T1 | T2 | T2 |
| Obs. | 1200 | 910 | 1260 | 940 |
| Cluster N | 20 | 20 | 21 | 21 |
| Controls | | V | | V |

Dyadic OLS regression. Dependent variable is a dummy which takes a value of one if i chose to establish a link with j. Each regression contains controls for a set of observables, round dummies and dummies for each possible pairing of map positions. Confidence: *** $\leftrightarrow$ 99%, ** $\leftrightarrow$ 95%, * $\leftrightarrow$ 90%. Standard errors are corrected for clustering at session level. P-values obtained with wild bootstrap-t procedure reported in parentheses.

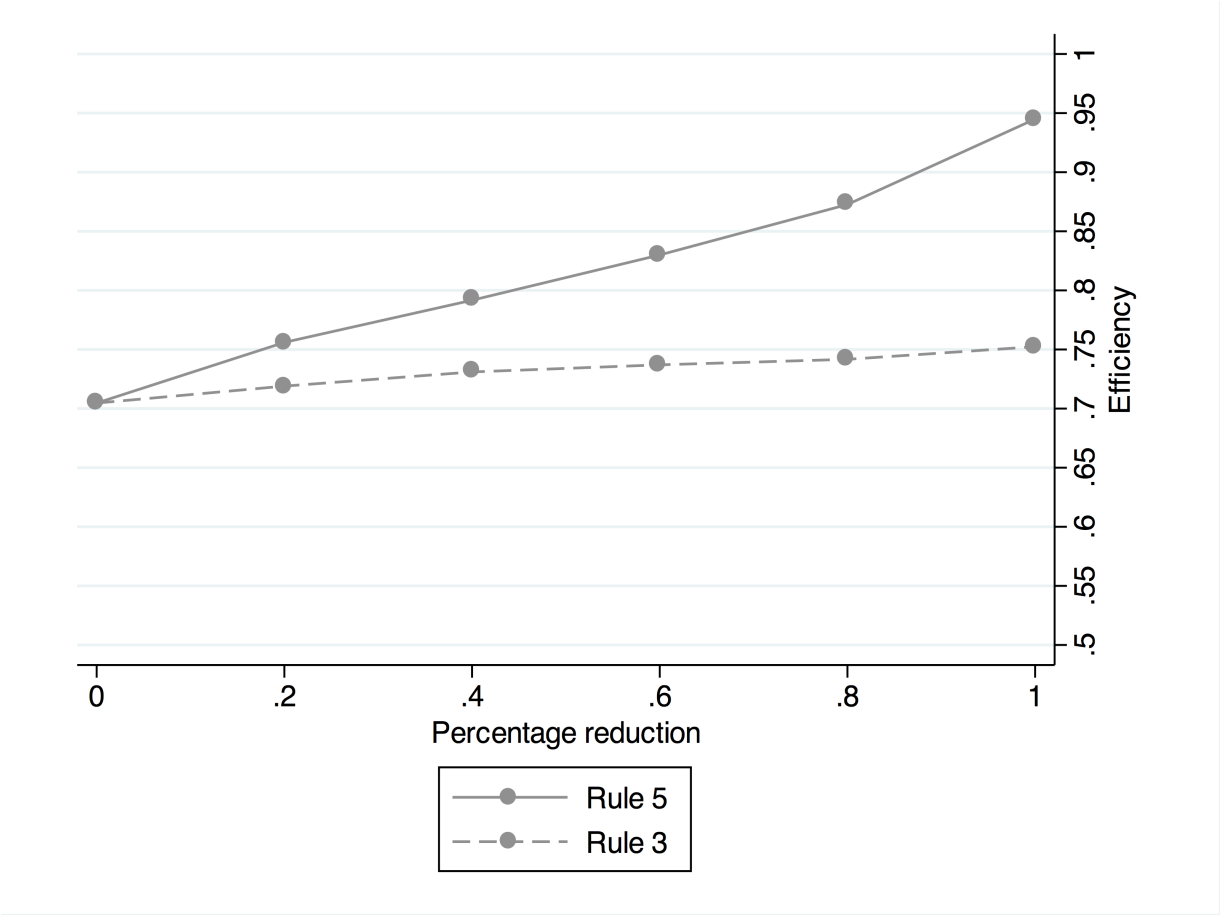
# Links to the most popular player



# Links to the most popular player

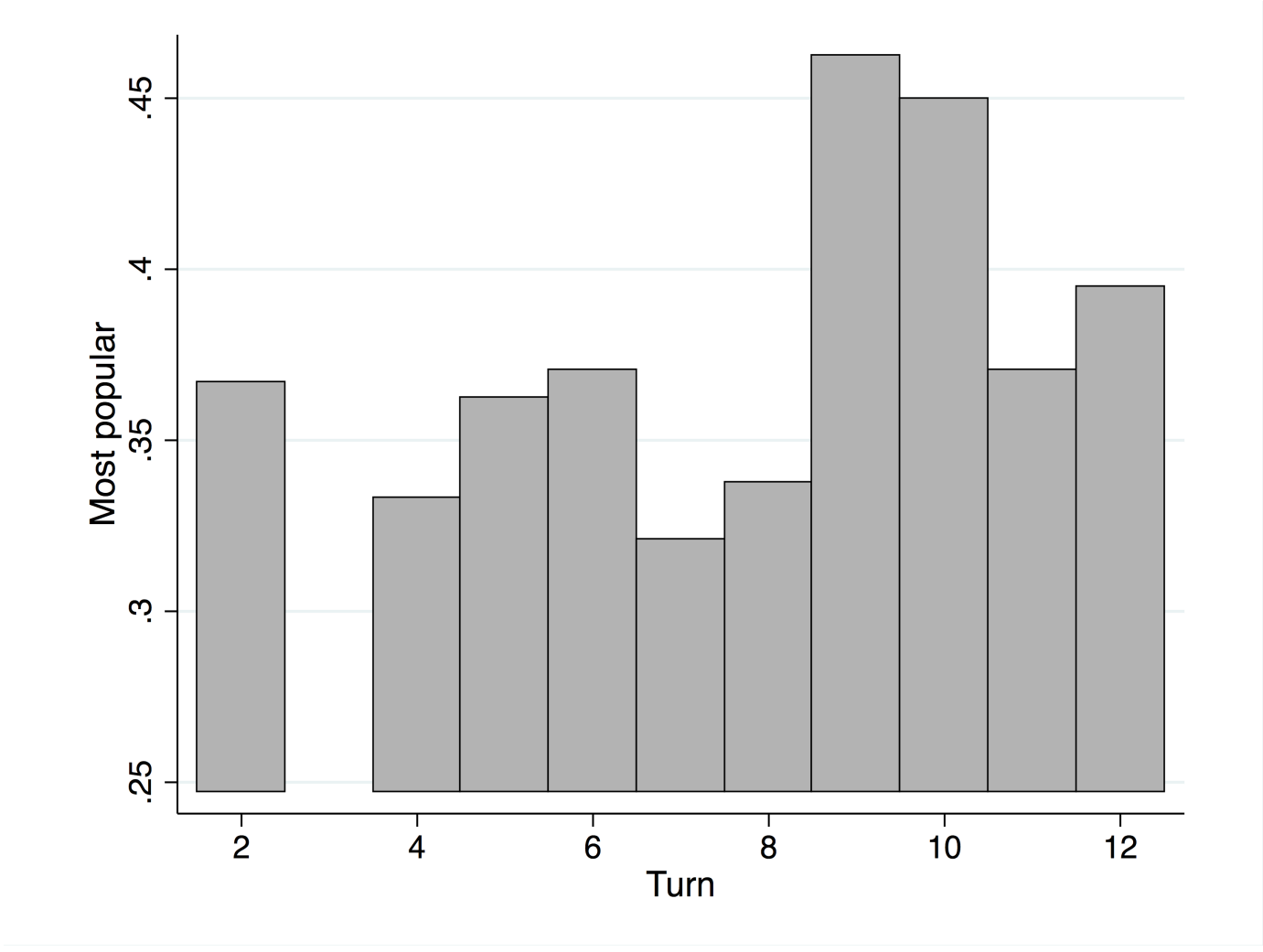


# **Result 3**: Rule 5 causes the highest efficiency losses



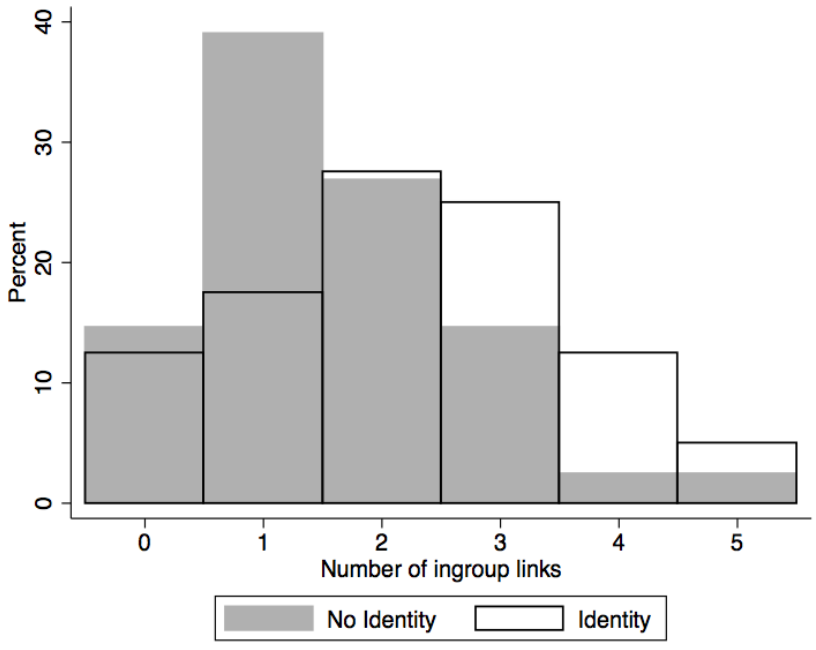

Note. In the baseline simulation 54 percent of decisions follow rule 2, 16 percent follow rule 3, and 30 percent follow rule 5. Each point in the graph represents average efficiency over 100 repetitions of the link formation game.

Figure : Proportion of links that target the 'most popular' player

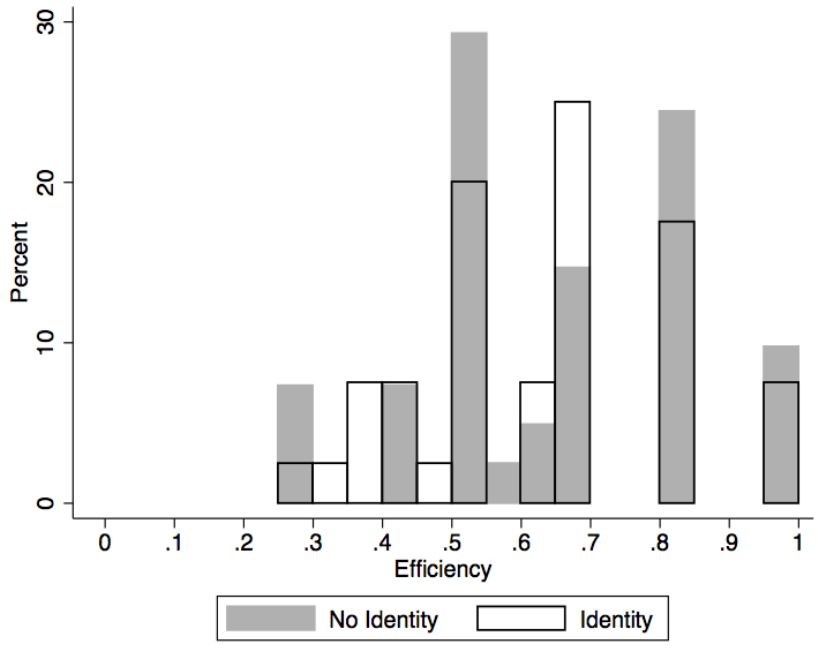


# **Result 4**: Ingroup links significantly increase, but there is no reduction in efficiency

Figure : Identity and no-identity sessions



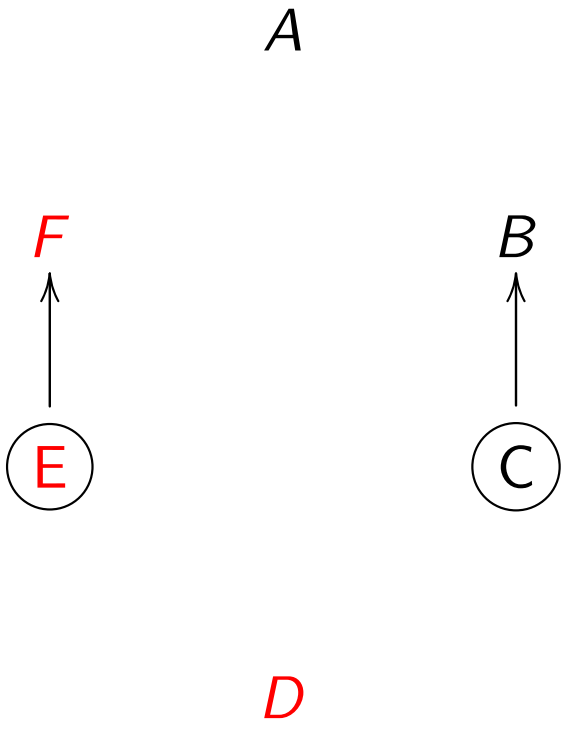

(a) **Ingroup links**
Z= 2.23, p= .02



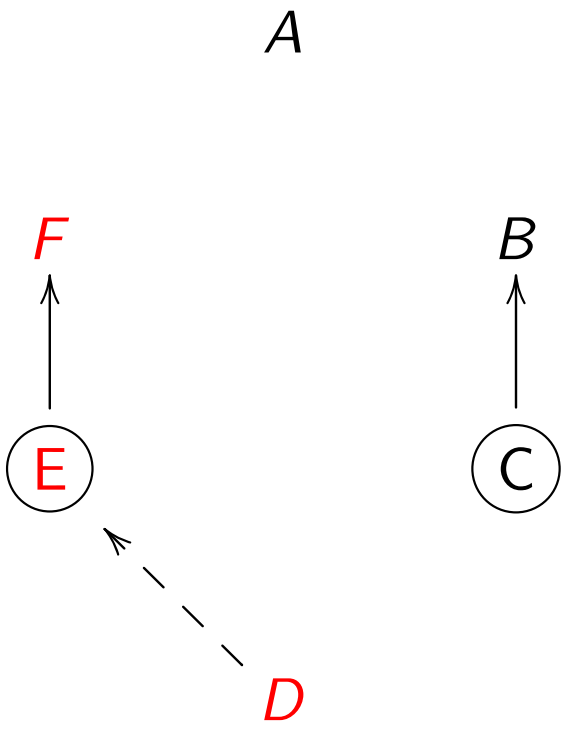

(b) **Efficiency**
Z= -0.51, p= .61

How does this come about?

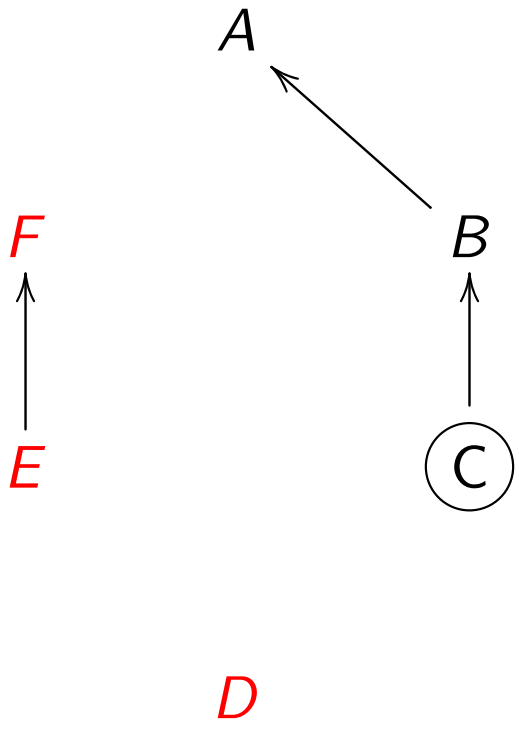
# (i) Best response set includes both in and out group players



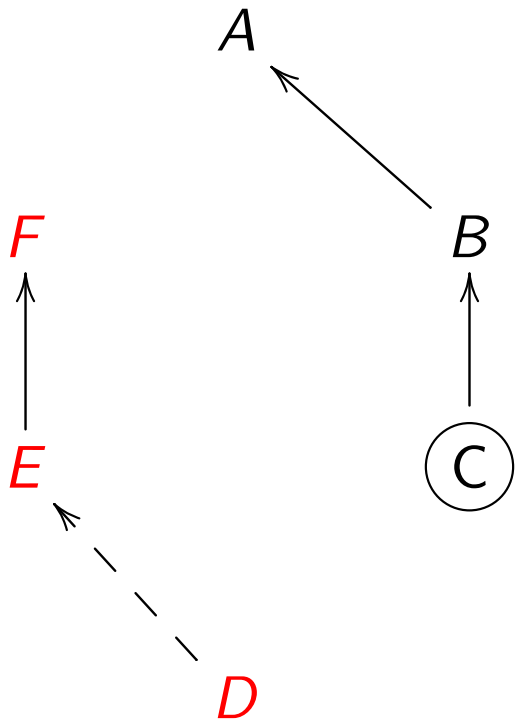
# Ingroup links ⇑, efficiency =



# (ii) Best response set includes only out group players



# Ingroup links ⇑, efficiency ⇓



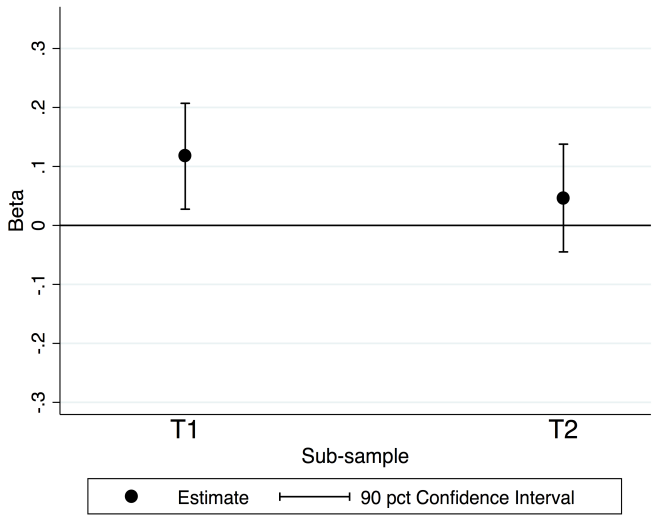
If group concerns are weak (these are arbitrary groups!), then farmers will choose in-group links only in case (i)

We will use the following model to explore mechanisms further:

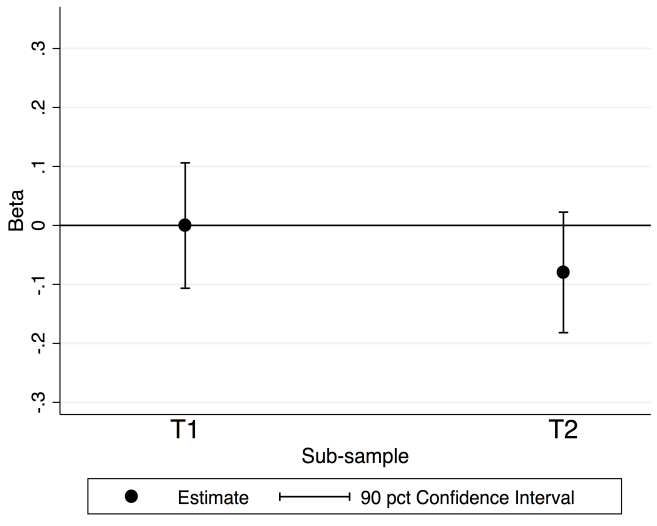
$$x_{dis} = \alpha + \beta_1 \text{Identity Session}_s + e_{dis} \quad (2)$$

# In-group links increase, efficiency minded links do not

Figure : Linear probability model (2): coefficient estimates

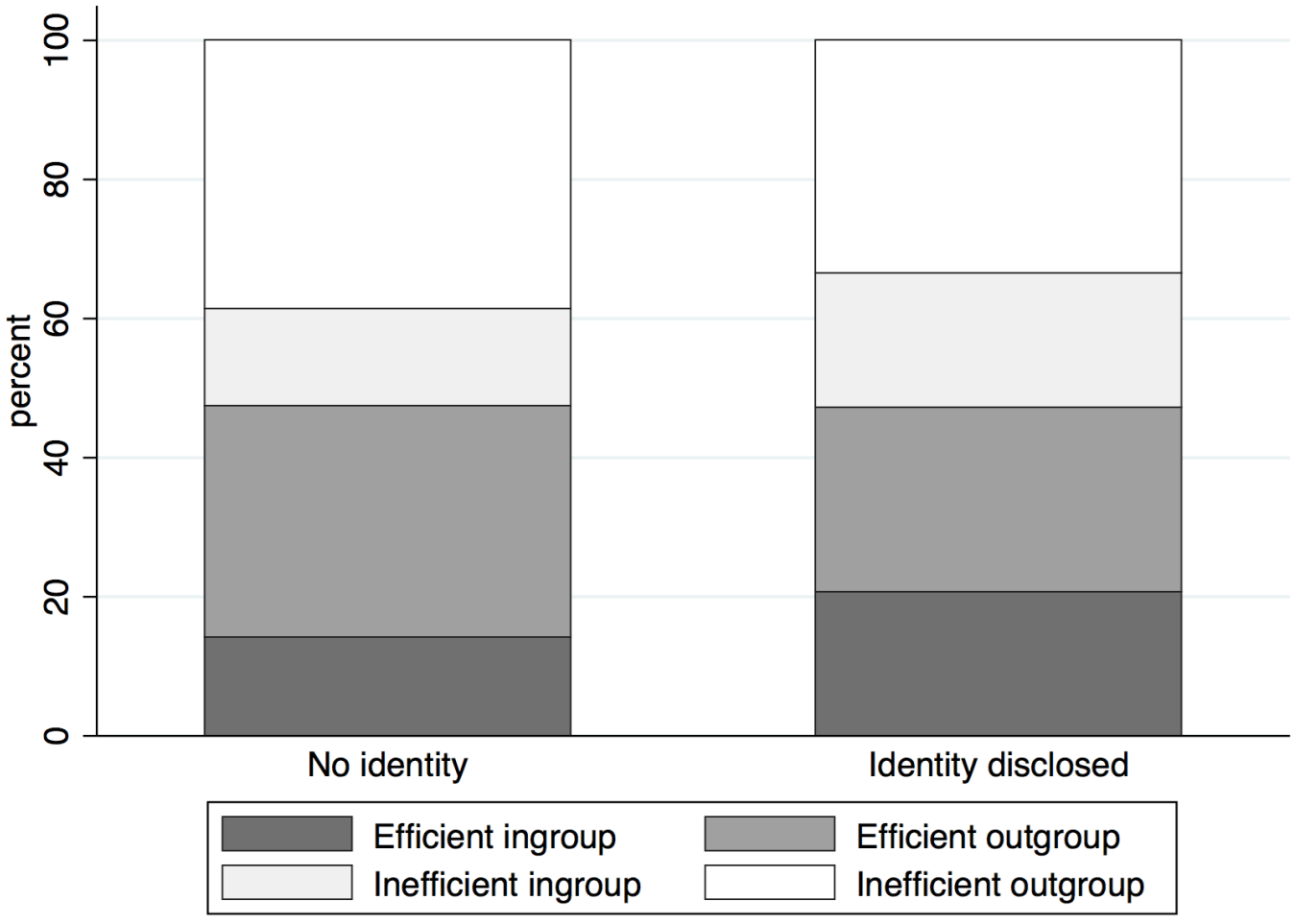



(a) $x_{dis}=$: in-group link




(b) $x_{dis}=$ efficiency-minded link

# Efficiency minded in-group links increase



# Outline

1 Design

2 Predictions

3 Results

4 Conclusion

Network efficiency is 32 percentage points below potential.

$\Rightarrow$ Future work needs to validate this with real networks and document the extent of belief convergence in communities

We show that group membership and a concern for other farmers' welfare affect link formation.

$\Rightarrow$ Important to understand information flows in society and the optimal design of peer-to-peer interventions

Network efficiency is 32 percentage points below potential.

$\Rightarrow$ Future work needs to validate this with real networks and document the extent of belief convergence in communities

We show that group membership and a concern for other farmers' welfare affect link formation.

$\Rightarrow$ Important to understand information flows in society and the optimal design of peer-to-peer interventions

# Balance

Table : Balance test: Identity Sessions

| | Age | Edu | UpperCaste | LandOwned | LandCult | NetSize |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Identity | -.194 | .029 | -.087 | .063 | .101 | -.201 |
| | (1.764) | (.056) | (.067) | (.517) | (.468) | (1.100) |
| Obs. | 479 | 466 | 433 | 475 | 470 | 428 |

OLS regressions. The dependent variable is indicated in the row's name. Upper caste is a variable that takes value of 1 if respondent is not from a schedule caste, a scheduled tribe or an Other Backward Caste. Network size is the self reported number of peers with whom the farmer exchanges advice on agricultural matters. Confidence: *** $\leftrightarrow$ 99%, ** $\leftrightarrow$ 95%, * $\leftrightarrow$ 90%. Standard errors clustered at the session level reported in parentheses.

# Balance

Table : Balance test: T2 sessions

| | Age | Edu | UpperCaste | LandOwned | LandCult | NetSize |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| T2 | -1.582 | -.028 | -.052 | -.085 | -.049 | 1.293 |
| | (1.761) | (.056) | (.068) | (.514) | (.465) | (1.089) |
| Obs. | 479 | 466 | 433 | 475 | 470 | 428 |

OLS regressions. The dependent variable is indicated in the row's name. Upper caste is a variable that takes value of 1 if respondent is not from a schedule caste, a scheduled tribe or an Other Backward Caste. Network size is the self reported number of peers with whom the farmer exchanges advice on agricultural matters. . Confidence: *** $\leftrightarrow$ 99%, ** $\leftrightarrow$ 95%, * $\leftrightarrow$ 90%. Standard errors clustered at the session level reported in parentheses.

# Balance

Table : Balance test: Identity in T1 sessions

| | Age | Edu | UpperCaste | LandOwned | LandCult | NetSize | Und |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Identity | -2.378 | .091 | -.040 | .077 | .141 | .111 | -.267 |
| | (2.544) | (.081) | (.098) | (.737) | (.657) | (1.102) | (.178) |
| Obs. | 235 | 232 | 215 | 234 | 231 | 211 | 240 |

OLS regressions. The dependent variable is indicated in the row's name. Upper caste is a variable that takes value of 1 if respondent is not from a schedule caste, a scheduled tribe or an Other Backward Caste. Network size is the self reported number of peers with whom the farmer exchanges advice on agricultural matters. Und is the number of mistakes in the initial 7 understanding questions. Confidence: *** $\leftrightarrow$ 99%, ** $\leftrightarrow$ 95%, * $\leftrightarrow$ 90%. Standard errors clustered at the session level reported in parentheses.

# Balance

Table : Balance test: Identity in T2 sessions

| | Age | Edu | UpperCaste | LandOwned | LandCult | NetSize | Und |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Identity | 1.879 | -.033 | -.135 | .046 | .061 | -.482 | -.224 |
| | (2.400) | (.076) | (.093) | (.733) | (.673) | (1.877) | (.197) |
| Obs. | 244 | 234 | 218 | 241 | 239 | 217 | 246 |

OLS regressions. The dependent variable is indicated in the row's name. Upper caste is a variable that takes value of 1 if respondent is not from a schedule caste, a scheduled tribe or an Other Backward Caste. Network size is the self reported number of peers with whom the farmer exchanges advice on agricultural matters. Und is the number of mistakes in the initial 8 understanding questions. Confidence: *** $\leftrightarrow$ 99%, ** $\leftrightarrow$ 95%, * $\leftrightarrow$ 90%. Standard errors clustered at the session level reported in parentheses.