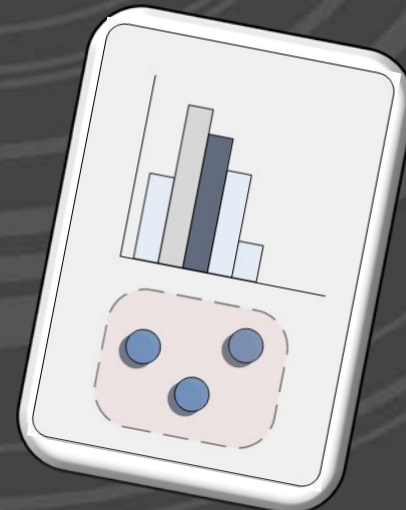
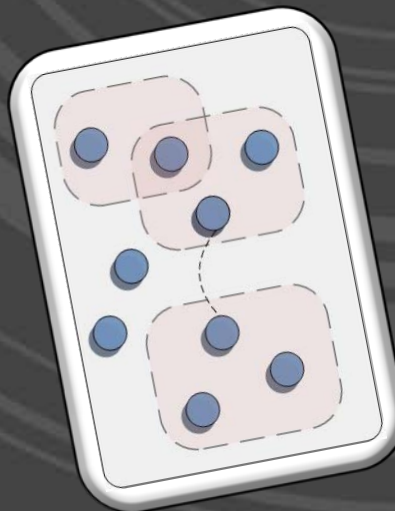


# Team Assembly and Scientific Collaboration on NanoHub

*Funded by National Science Foundation Grant # IIS-299269*



**Drew Margolin**  
**Katherine Ognyanova**  
University of Southern California



**Meikuan Huang**  
**Yun Huang**  
**Noshir Contractor**  
Northwestern University

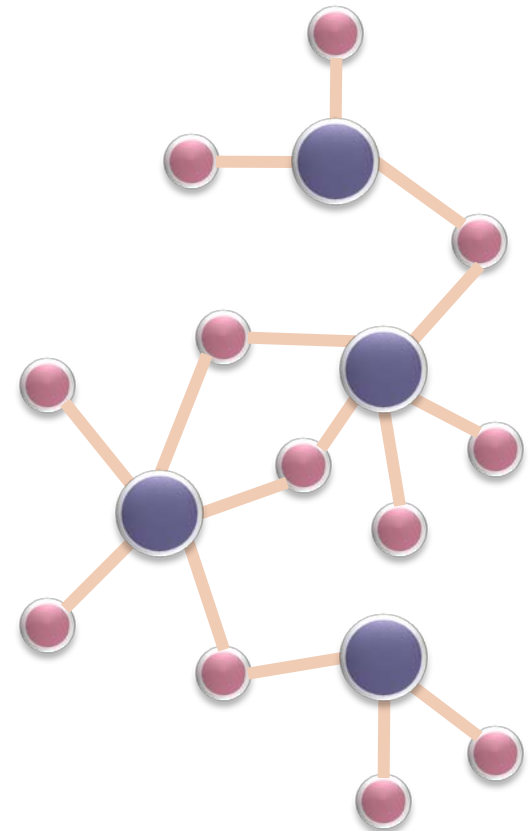
## **Team Formation and Performance on Nanohub: A Network Selection Challenge in Scientific Communities**

**Goal:** Submission to "Networks in Social Policy Problems",  
Edited Book, Cambridge University Press

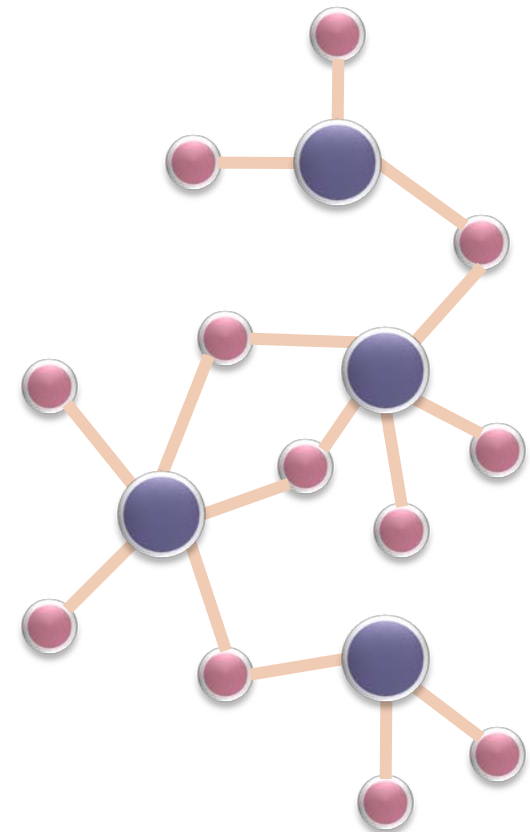
# Team Assembly – Theoretical Setup

✓ Team definition: *voluntary collaborative project teams* (c.f. Schwab & Miner, 2008)

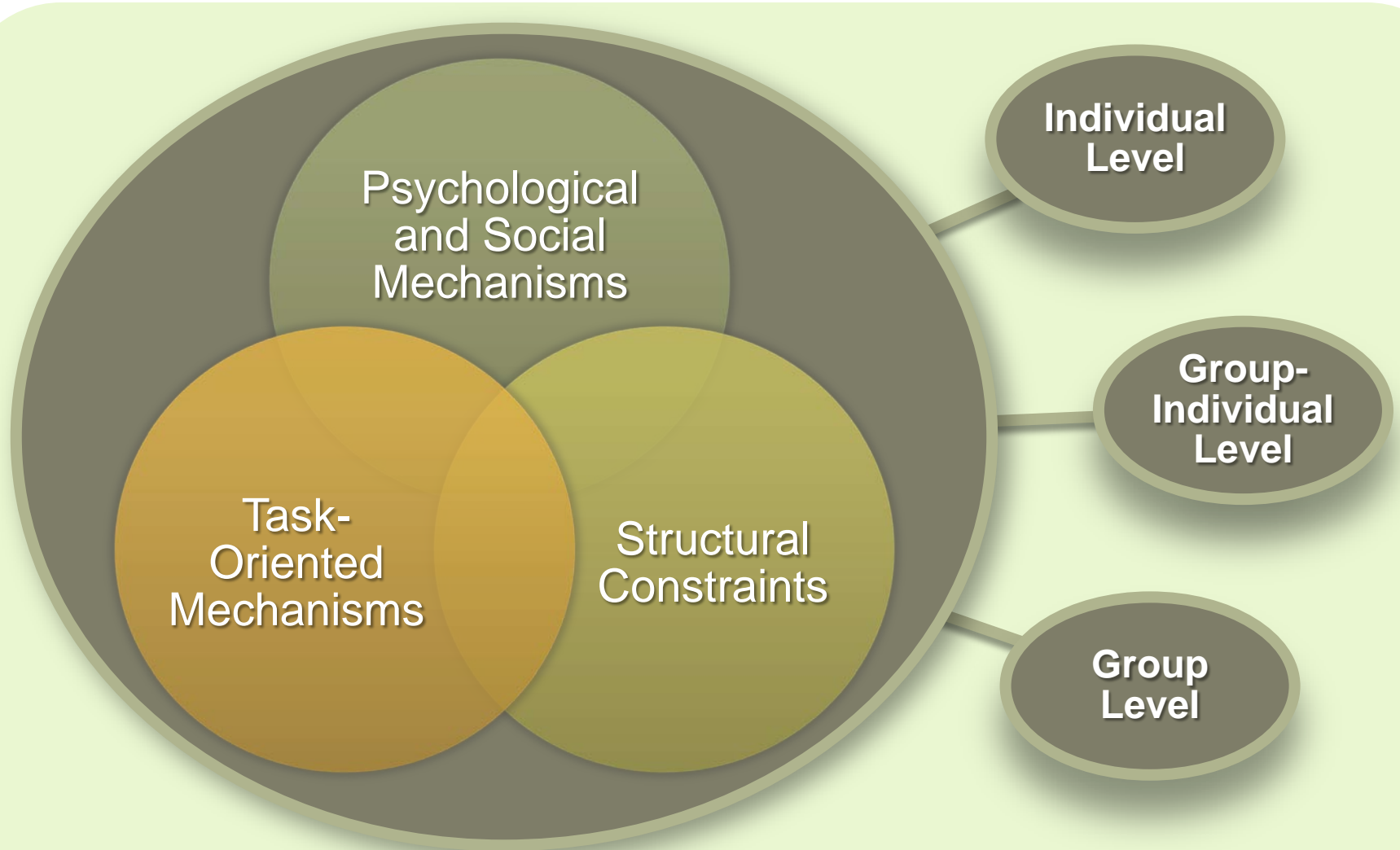
- The members are not appointed by an authority or institution
- The team is goal-oriented: assembled explicitly for the purpose of completing a task
- In theory, VCPTs should be assembled with view of optimizing for high performance



- ✓ **Top-Down Guidance:** Accessing local knowledge at the team level is usually difficult for managers and policy-makers. Policy is often targeted towards general best practices rather than one tailored to specific teams and/or communities.
- ✓ **Bottom-Up Emergence:** Conformity and aspirations for legitimacy affect the rules and norms developed by team members. We see attempts for imitation of successful models, adherence to institutional pressures, and other factors that may not be beneficial for team performance.



# Selection Mechanisms and Community Logics in Team Assembly



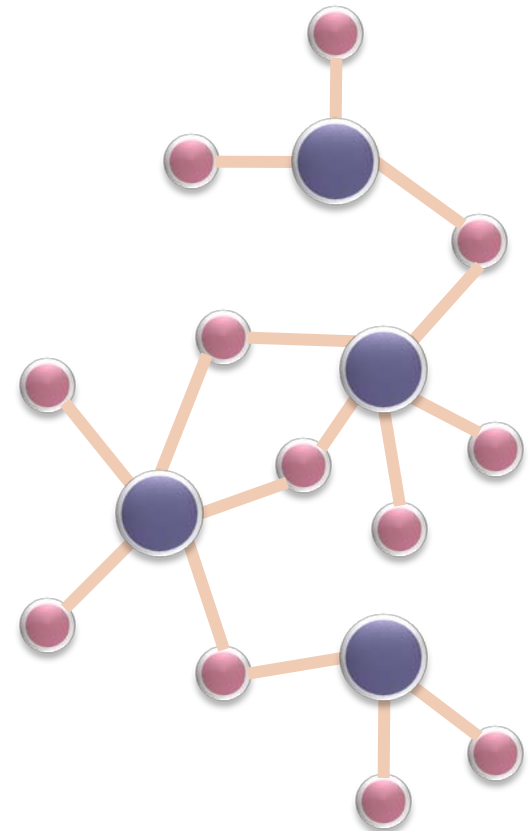
# Network Analysis: Signatures of Community Norms

*Network analysis can help us identify emerging community logics of team assembly. Once we know the logics, we can analyze their effect on performance.*

**H1:** For some logics of team assembly that are typical in a community, conformity to this logic will be associated with positive performance.

**H2:** For some emergent logics of team assembly that are typical in a community, conformity to this logic will be associated with negative performance.

**RQ1:** Are any atypical, experimental logics of team assembly performance enhancing?



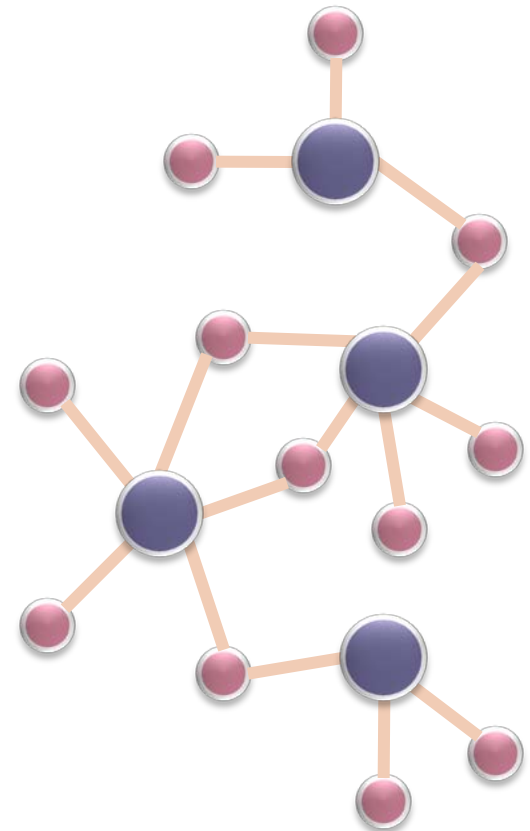
## 1. Inferring community logics: ERGM

- Identify dominant tendencies in team formation
- Model parameters based on existing literature

## 2. Assessing impact on performance: Regression

- Measures representing the community logics in (1) are regressed on a performance measure

## 3. *Heckman correction* – typical vs. atypical logics identified through simulations (Not yet implemented)



## Bipartite Network

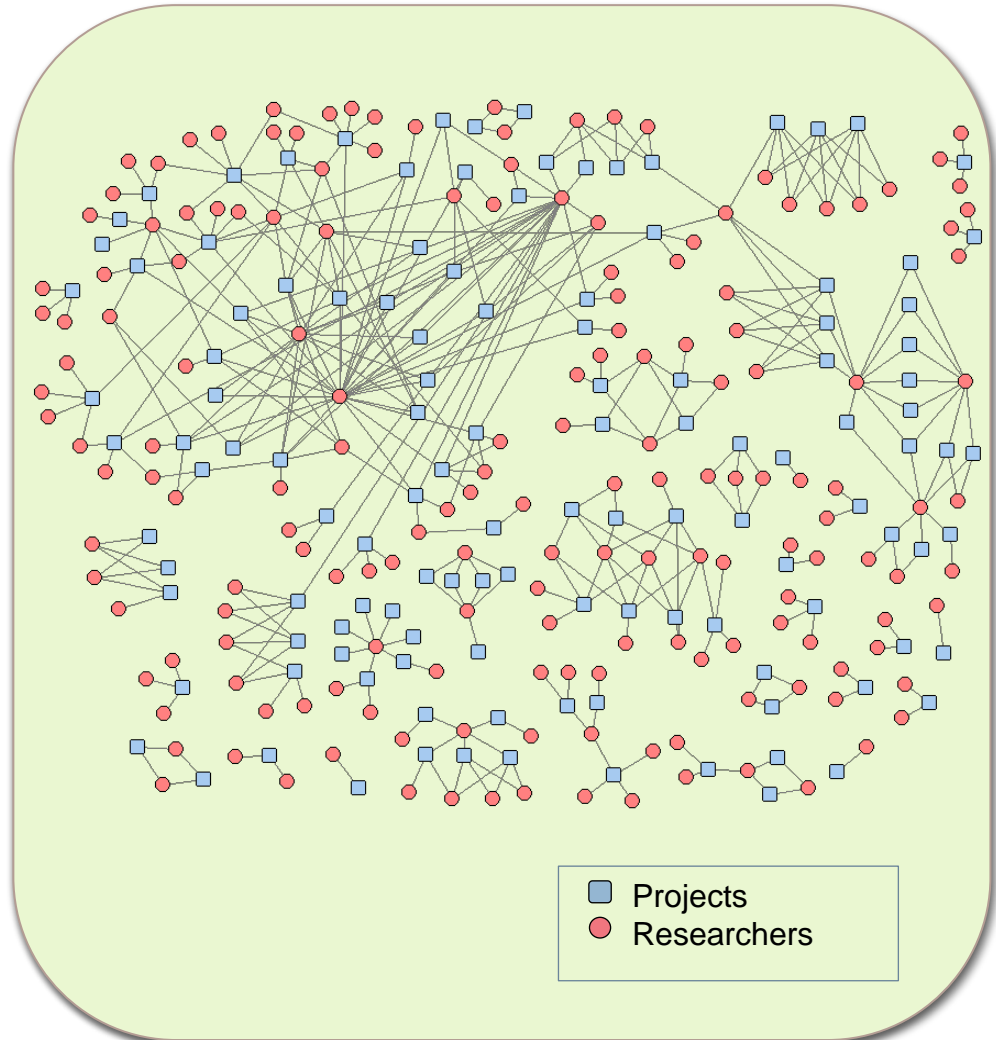
- **170 Researchers**
  - 161 male : 9 female
  - Coming from 9 countries
- **124 Teams**
  - Sizes vary from 1 to 8

## Projects

- **124 Software tools**  
(education, computation, simulations)
- **Difficulty level (1-4)**  
(easy, intermediate, advanced, expert)
- **Difficult projects (level 3-4)**

## Expertise

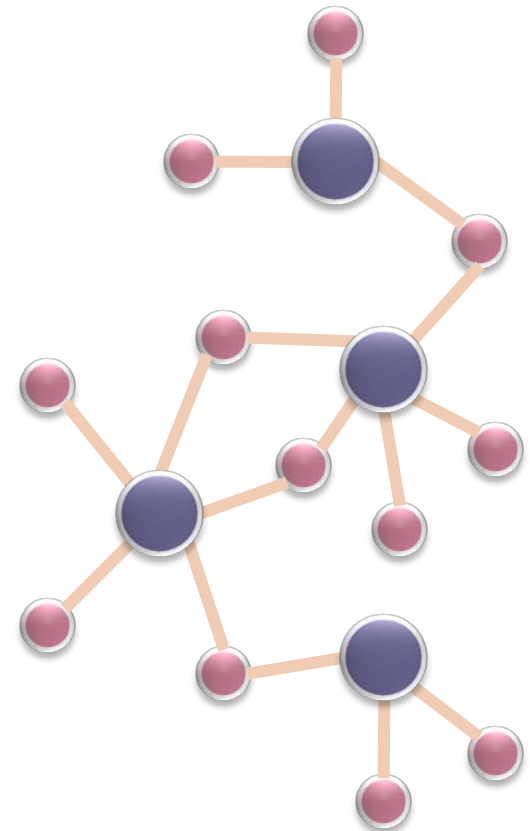
- **Number of publications**  
(Web of Science)
- **Experts: top 25% of researchers**  
(N=43)



## Five underlying community logics tested:

1. **Gender-related:** preference for selecting females
2. **Expertise-driven:** preference for selecting experts
3. **NCN-driven:** tendency to include NCN members (this is a nanoHub specific logic)
4. **Star-related:** the propensity of a few popular or extraverted individuals to participate in a large number teams (star =  $\text{degree} > M + 1SD$ )
5. **Team size:** included as a control variable as it has a complex role in patterns of team assembly and performance.

*Controls: team stars, alternating k-2-paths, density*



# Step 2: Regression

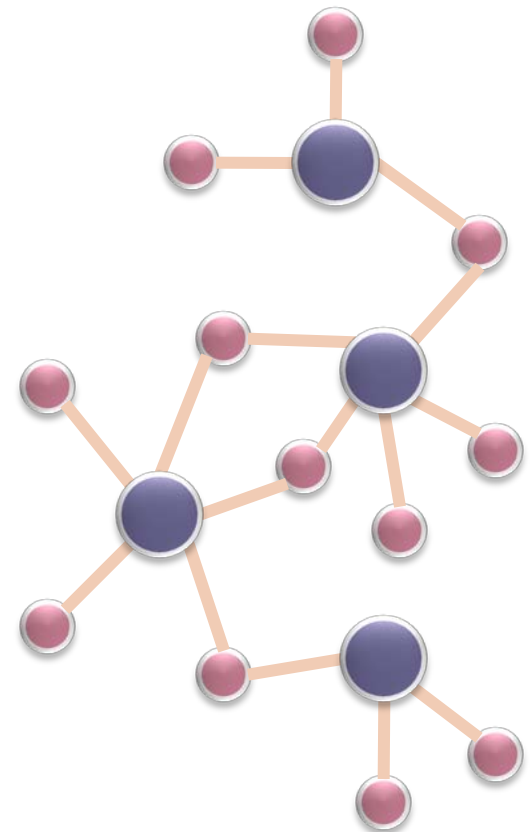
✓ **Independent variables**

Same as ERGM parameters. For each team: size, number of females, experts, NCN members, stars.

✓ **Dependent Variable**

Performance: composite measure based on five indicators (computed as sum of z-scores)

- Tool Rating
- Number of Times Rated
- Number of Times Cited
- Number of Times Tagged
- Number of Taggers



Effects	Estimates	SE	t-Ratio	Significance
Density	-2.737	0.476	-0.003	*
Member stars	0.162	0.268	-0.010	-
Team stars	-0.860	0.275	0.004	*
Concurrent collaboration	-0.205	0.118	-0.026	-
Experts on teams	0.107	0.133	0.008	-
Females on teams	0.177	0.267	0.053	-
NCN members on teams	4.903	1.471	-0.035	*

# Results: Regression

Table 2

*Regression model: Community logics - impact on performance.*

Predictors	Standardized Regression Coefficients ( $\beta$ )	Zero-order correlation with performance	Partial correlation with performance
Team Size	.176	.397*	.161
Number of Females	.183*	.296*	.182*
Number of Experts	.236*	.401*	.207*
Number of Stars	.424*	.341*	.260*
Number of NCN	-.457*	.165*	-.286*

\* $p < .05$

$R^2 = .29$ , adjusted  $R^2 = .26$ ,  $F(5, 118) = 9.54$ ,  $p < 0.001$

Table 3  
*Correlations between predictor variables.*

Predictors	Team Size	Number of Females	Number of Experts	Number of Stars	Number of NCN
Team Size	1	.302*	.569**	.494**	.389**
Number of Females		1	.385**	.452**	.486**
Number of Experts			1	.519**	.494**
Number of Stars				1	.821**
Number of NCN					1

\* $p < .05$  \*\* $p < 0.001$

# Results: Hypothesis Testing

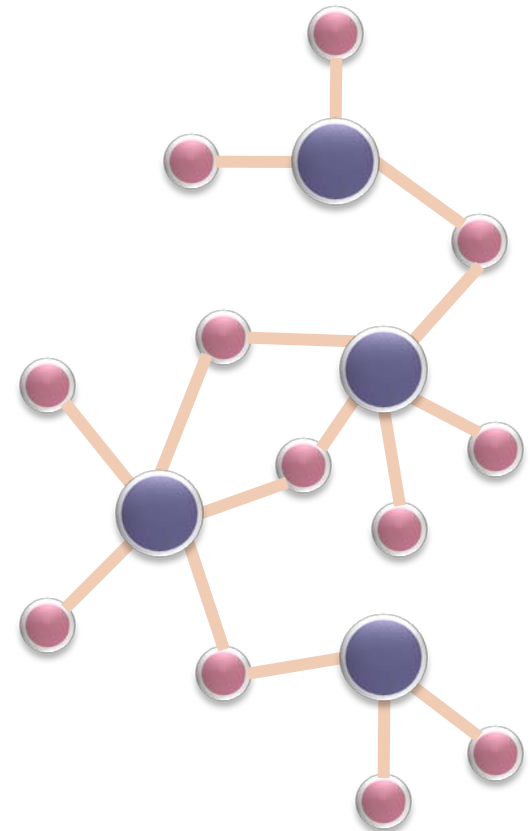
<b>Team Attribute</b>	<b>Typical Community logic?</b> (Based on ERG model)	<b>Effect on Performance?</b> (Based on regression model)	<b>Hypothesis Testing</b> (logics and outcomes)
Team size	<b>( Control variable)</b>	<b>N/A</b>	<b>N/A</b>
Gender	<b>NO</b>	<b>POSITIVE</b>	<b>RQ1 - Tentative Yes</b>
Experts	<b>NO</b>	<b>POSITIVE</b>	<b>RQ1 - Tentative Yes</b>
Stars	<b>NO</b>	<b>POSITIVE</b>	<b>RQ1 - Tentative Yes</b>
NCN members	<b>YES</b>	<b>NEGATIVE</b>	<b>Supports H2</b>

## ✓ Potential policy suggestions:

- Set a quota for NCN members
- Provide more opportunities for collaboration with experts and stars
- Encourage more females to join platform

## ✓ Next steps to enhance the study:

- A more comprehensive ERG model including other relevant variables (disciplines, diversity, etc.)
- Implementing a network form of the Heckman correction using simulations



**Thank you!**