Specific and General Information Sharing Among Academic Scientists

Carolin Haeussler
Ludwig Maximillian University

Jerry Thursby
Georgia Tech

Marie Thursby
Georgia Tech & NBER

2nd Communia Conference, Turin, Italy
Widely recognized that sharing of research results is essential for scientific progress.

- Mertonian norm of communalism in principal enforced by priority based reward system (Stephan 96)

- Winner-take-all yields tension between communal needs & competitive pressures (Dasgupta & David 94, Murray & O’Mahoney 07)

Growing attention to communication among researchers both theoretically and empirically

- Theory: Stein 08, Anton and Yao 02/04, Hellmann and Perotti 07, Lerner & Tirole 02, Gill 08, Mukherjee & Stern 09, Lacetera and Zirulia 08

- Empirics: Blumenthal et al. 1996, Campbell et al. 2000, Walsh et al. 07, Haeussler 09
We examine what drives scientist decisions to share or not.

We consider 2 types of sharing:

- Specific sharing (one on one sharing)
- General sharing (present to audiences)

For each we develop a game theoretic models and use survey evidence to test implications of the models.
Specific Sharing Game

Two scientists trying to solve problem with prize $W$

- Choosing whether to share data or materials with each other

Simultaneous move game—a Prisoner’s Dilemma

- 1’s Pr of $W$ is $z$ and 2’s is $(1-z)$ if they make same choice
- If either unilaterally shares, her pr of $W$ decreases
  \[ q < z \text{ and } (1-q) > z \]
- Scientist $i$ has $r_i$ to share
- Ability of $j$ to exploit $i$’s data is $e_j$
- Unique Nash Eqm is (NS, NS) but (S,S) Pareto dominates

<table>
<thead>
<tr>
<th>Scientist 1</th>
<th>SHARE</th>
<th>NOT SHARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHARE</td>
<td>$zW + e_1r_2$</td>
<td>$qW$</td>
</tr>
<tr>
<td></td>
<td>$(1-z)W + e_2r_1$</td>
<td>$(1 - q)W + e_2r_2$</td>
</tr>
<tr>
<td>NOT SHARE</td>
<td>$(1-q)W + e_1r_2$</td>
<td>$zW$</td>
</tr>
<tr>
<td></td>
<td>$qW$</td>
<td>$(1 - z)W$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scientist 2</th>
<th>SHARE</th>
<th>NOT SHARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHARE</td>
<td>$zW + e_1r_2$</td>
<td>$qW$</td>
</tr>
<tr>
<td></td>
<td>$(1-z)W + e_2r_1$</td>
<td>$(1 - q)W + e_2r_2$</td>
</tr>
<tr>
<td>NOT SHARE</td>
<td>$(1-q)W + e_1r_2$</td>
<td>$zW$</td>
</tr>
<tr>
<td></td>
<td>$qW$</td>
<td>$(1 - z)W$</td>
</tr>
</tbody>
</table>
Repeated Specific Sharing Game

Trigger strategies
• Each shares as long as the other does
• Refuses in subsequent periods once the other does not

Sharing for some time can be sgp eqm in probabilistic horizon game
• Sufficiently long time of expected play \( \alpha/(1-\alpha) \)
• Low enough period gain from NS today (depends on \( z, q, \) and \( W \))
• High enough period loss from NS tomorrow (depends on \( e_j \) and \( r_i \))

Likelihood of sharing in eqm
• Decreases the higher the value of the prize, \( W \)
• Increases with the expected length of the game, \( q, e_j \) and \( r_i \)
• More similar the scientists the more likely they are to share
General Sharing Game

$M \geq 2$ scientists trying to solve problem with prize $W$

- Scientist 1 has solved a portion of the problem ($\sigma$)
- No one has totally solved
- Pr that a randomly chosen scientist has solved a different part is $\gamma$
- $\Lambda = 1 - (1 - \gamma)^{(M-1)}$ at least one scientist is a complementor

Scientist 1 considers whether to make $\sigma$ public

- Positive effects
  - Announce her progress for credit $\sigma W$
  - Feedback from complementors
- Negative effects
  - Increases Pr a complementor wins (reduces 1’s Pr to $x-\delta$)
  - Risk that a winner will not acknowledge her part of the solution
General Sharing Game (continued)

Structure of game

- Scientist 1 decides to present (P) or not (NP)
- After presentation all continue working on problem
- Nature picks winner
- Winner chooses to acknowledge (A) or not (NA)

All scientists have belief $\rho$ that a randomly chosen scientist will verify whether $\sigma$ is acknowledged

- Belief of verification: $v = 1 - (1 - \rho)^{(M-2)}$
- If NA is verified, winner loses reputation $R$
General Sharing Game (continued)

A dominates NA if \( v > \sigma W/(R+W) \)

Likelihood of A in equilibrium

- Increases with
  - Size of the community (M)
  - Belief that a randomly chosen scientist will verify (\( \rho \))
  - Reputation loss (R)
- Decreases with the portion of the problem scientist 1 has solved (\( \sigma \))

Scientist 1 chooses to present if

- Expected utility of P > NP
- Given \( C = \Pr(A) + v\Pr(NA) \)
General Sharing Game (continued)

• Likelihood of P in equilibrium
  Increases with
    • Feedback (τ)
    • Belief a randomly chosen scientist will verify (ρ)
    • Reputation loss (R)
  Decreases with
    • Prize (W)
    • Pr 1 wins without presenting (x)
    • Improvement to complementor’s Pr of W
  Increases with
    • Size of the community (M) when τ > δW
Survey and Data

- Surveyed academic and industrial bio-scientists in 2007
  - Sample: British and German scientists that
    - filed at least one patent at the European Patent Office in the biotechnical area and/or
    - Published at least one article listed in PubMed in the biotechnical area.
  - Final sample: 2452 scientists identified in the EPO-database; 2169 scientists identified in PubMed (33% and 23%). 1087 British and 3067 German respondents.
Survey and Data

• For this study
  – Only British and German public sector scientists
  – Excluded scientists older than 65 years
  – Final sample: 1173 academic scientist (21% from United Kingdom)
<table>
<thead>
<tr>
<th>Question</th>
<th>Type of Sharing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 I present unpublished or yet to be patented research results at conferences.</td>
<td>General</td>
</tr>
<tr>
<td>2 When I discuss unpublished or yet to be patented research results, I often withhold crucial parts</td>
<td>General</td>
</tr>
<tr>
<td>3 In the past I have delayed or had to delay publication of my research in order to secure patenting the research results.</td>
<td>General</td>
</tr>
<tr>
<td>4 I only discuss unpublished or yet to be patented research results with people who will for sure not pass on this information.</td>
<td>Specific</td>
</tr>
<tr>
<td>5 I only discuss unpublished or yet to be patented research results with people from whom I expect valuable feedback.</td>
<td>Specific</td>
</tr>
<tr>
<td>6 Before I share unpublished or yet to be patented research results, I first consider whether or not I will get valuable information from this researcher in the future.</td>
<td>Specific</td>
</tr>
</tbody>
</table>
Econometric Model

Separate models for specific and general questions

• Dependent variable 5-point Likert Scale.
• Larger values = greater willingness to share.
• The 3 specific questions are used in a panel.
• The 3 general questions are used in a panel.
• Ordered logit with cluster standard errors.
Independent Variables

*Competition* Level of competition (1-5)
*FirstEsteemed* First to discover is esteemed (1-5)
*TeamSize* Size of research team
*Age* Age of respondent
*AgeSq* Age squared
*Professor* = 1 if respondent is professor
*Responsible* Number who report directly to respondent
*OpenExchange* Open exchange is usually practiced (1-5)
*Publications* Number of publications
*Basic* Work is basic (1-5)
*OwnResearch* % of effort on own research
*Patents* Number of patents
*Consult* % of effort on consulting
*FamilyEnt* = 1 if family member is entrepreneur
*Married* = 1 of married
*Male* = 1 if male
*UK* = 1 if UK scientist
*ExploitLose* If exploit work of others, then lose (1-5)

Question Fixed Effects
Field fixed effects
<table>
<thead>
<tr>
<th>Variable</th>
<th>Odds Ratio</th>
<th>t-Stat</th>
<th>Model Prediction</th>
<th>Our Prior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competition</td>
<td>0.8850</td>
<td>-2.54**</td>
<td>Dec</td>
<td></td>
</tr>
<tr>
<td>FirstEsteemed</td>
<td>0.9088</td>
<td>-1.82*</td>
<td>Dec</td>
<td></td>
</tr>
<tr>
<td>TeamSize</td>
<td>1.0054</td>
<td>2.52**</td>
<td>Inc</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>1.0842</td>
<td>1.34</td>
<td>Dec</td>
<td></td>
</tr>
<tr>
<td>AgeSq</td>
<td>0.9992</td>
<td>-1.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professor</td>
<td>1.3663</td>
<td>2.92***</td>
<td>Inc</td>
<td></td>
</tr>
<tr>
<td>Responsible</td>
<td>0.9981</td>
<td>-1.33</td>
<td>Dec</td>
<td></td>
</tr>
<tr>
<td>OpenExchange</td>
<td>1.1988</td>
<td>3.09***</td>
<td>Inc</td>
<td></td>
</tr>
<tr>
<td>Publications</td>
<td>1.0002</td>
<td>0.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic</td>
<td>1.1268</td>
<td>2.36**</td>
<td>Inc</td>
<td></td>
</tr>
<tr>
<td>OwnResearch</td>
<td>0.9979</td>
<td>-0.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patents</td>
<td>0.9887</td>
<td>-2.26**</td>
<td>Dec</td>
<td></td>
</tr>
<tr>
<td>Consult</td>
<td>0.9925</td>
<td>-1.05</td>
<td>Dec</td>
<td></td>
</tr>
<tr>
<td>FamilyEnt</td>
<td>1.0570</td>
<td>0.53</td>
<td>Dec</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>1.2357</td>
<td>1.67*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.9395</td>
<td>-0.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>1.1769</td>
<td>1.37</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Question & field fixed effects

r-Square          0.0360
Obs.              3103

*** Significant at 1%   ** Significant at 5%   * Significant at 10%
## General Sharing

<table>
<thead>
<tr>
<th>Variable</th>
<th>Odds Ratio</th>
<th>t-Stat</th>
<th>Model Prediction</th>
<th>Our Prior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competition</td>
<td>0.8887</td>
<td>-2.50**</td>
<td>Ambiguous</td>
<td>Dec</td>
</tr>
<tr>
<td>FirstEsteemed</td>
<td>0.9952</td>
<td>-0.10</td>
<td>Dec</td>
<td></td>
</tr>
<tr>
<td>Teamsize</td>
<td>0.9937</td>
<td>-1.52</td>
<td>Dec</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.9855</td>
<td>-0.26</td>
<td>Inc</td>
<td></td>
</tr>
<tr>
<td>AgeSq</td>
<td>1.0000</td>
<td>0.00</td>
<td>Inc</td>
<td></td>
</tr>
<tr>
<td>Professor</td>
<td>0.9761</td>
<td>-0.25</td>
<td>Dec</td>
<td></td>
</tr>
<tr>
<td>Responsible</td>
<td>0.9953</td>
<td>-3.73***</td>
<td>Dec</td>
<td></td>
</tr>
<tr>
<td>OpenExchange</td>
<td>1.3048</td>
<td>5.30***</td>
<td>Inc</td>
<td></td>
</tr>
<tr>
<td>Publications</td>
<td>1.0013</td>
<td>1.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic</td>
<td>1.1990</td>
<td>3.92***</td>
<td>Inc</td>
<td></td>
</tr>
<tr>
<td>OwnResearch</td>
<td>1.0018</td>
<td>0.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patents</td>
<td>0.9468</td>
<td>-2.30**</td>
<td>Dec</td>
<td></td>
</tr>
<tr>
<td>Consult</td>
<td>0.9770</td>
<td>-2.60***</td>
<td>Dec</td>
<td></td>
</tr>
<tr>
<td>FamilyEnt</td>
<td>0.8251</td>
<td>-1.98**</td>
<td>Dec</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>1.0900</td>
<td>0.84</td>
<td>Dec</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.7389</td>
<td>-3.06***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>0.8777</td>
<td>-1.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ExploitLose</td>
<td>0.9270</td>
<td>-1.85*</td>
<td>Inc</td>
<td></td>
</tr>
</tbody>
</table>

Question & field fixed effects

r-Square: 0.0706
Obs: 3063

### Significance Levels
- **Significant at 1%**
- **Significant at 5%**
- *Significant at 10%*
Willingness to Share: General vs. Specific

![Bar chart showing comparison between willingness to share general and specific information. The x-axis represents the willingness to share (from low to high), and the y-axis represents the probability of sharing. The chart indicates higher willingness to share specific information compared to general information across different levels of willingness. ]
Concluding Remarks

A simple theory to frame empirical analysis of sharing

• Specific and general sharing—except for the impact of competition—should be viewed differently
  • more funding attracts more scientists but eases competition for funding.

• A question for future work is to endogenize M so that the impact of research funding can be better understood
Concluding Remarks

Acknowledgment of intermediate results important for general sharing

• Strong implications for journal policies which are hardly strong (Lacetera and Zirulia 2008, Enders and Hoover 2004)

• Exception (*Nature* 2009)

• Technological advances help (*Science* May 2009, Couzin-Frankel and Grom 2009)