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 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**

		Scientist 2	
		SHARE	NOT SHARE
Scientist 1	SHARE	$zW + e_1 r_2$ $(1-z))W + e_2 r_1$	qW $(1-q)W + e_2 r_2$
	NOT SHARE	$(1-q)W + e_1 r_2$ qW	zW $(1-z)W$

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≥2 scientists trying to solve problem with prize W

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

Scientist 1 considers whether to make σ public

- Positive effects
  - Announce her progress for credit σW
  - Feedback from complementors
- Negative effects
  - Increases Pr a complementor wins (reduces 1's Pr to x-δ)
  - 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 ρ that a randomly chosen scientist will verify whether σ 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 (ρ)**
  - **Reputation loss (R)**
- **Decreases with the portion of the problem scientist 1 has solved (σ)**

Scientist 1 chooses to present if

- **Expected utility of P > NP**
- **Given C = Pr(A) + vPr(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 $\tau > \delta W$**

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 1. Sharing Questions

	Question	Type of Sharing
1	I present unpublished or yet to be patented research results at conferences.	General
2	When I discuss unpublished or yet to be patented research results, I often withhold crucial parts	General
3	In the past I have delayed or had to delay publication of my research in order to secure patenting the research results.	General
4	I only discuss unpublished or yet to be patented research results with people who will for sure not pass on this information.	Specific
5	I only discuss unpublished or yet to be patented research results with people from whom I expect valuable feedback.	Specific
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.	Specific

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

Specific Sharing

Variable	Odds Ratio	t-Stat	Model Prediction	Our Prior
Competition	0.8850	-2.54**	Dec	
FirstEsteemed	0.9088	-1.82*	Dec	
TeamSize	1.0054	2.52**	Inc	
Age	1.0842	1.34	Dec	
AgeSq	0.9992	-1.29		
Professor	1.3663	2.92***	Inc	
Responsible	0.9981	-1.33		Dec
OpenExchange	1.1988	3.09***		Inc
Publications	1.0002	0.23		
Basic	1.1268	2.36**		Inc
OwnResearch	0.9979	-0.71		
Patents	0.9887	-2.26**		Dec
Consult	0.9925	-1.05		Dec
FamilyEnt	1.0570	0.53		Dec
Married	1.2357	1.67*		
Male	0.9395	-0.54		
UK	1.1769	1.37		
Question & field fixed effects				
r-Square	0.0360			
Obs.	3103			

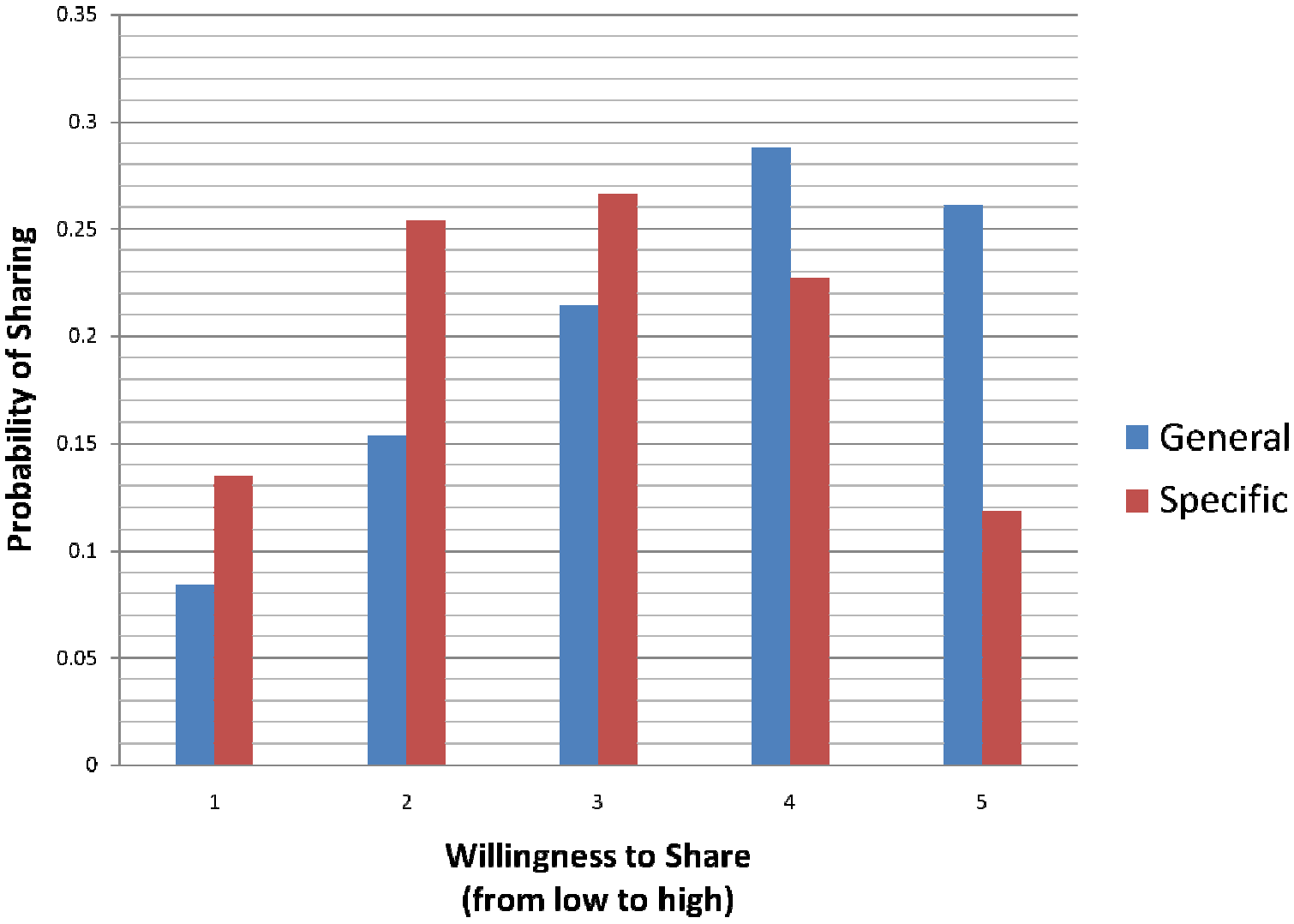
*** Significant at 1% ** Significant at 5% * Significant at 10%

General Sharing

Variable	Odds Ratio	t-Stat	Model Prediction	Our Prior
Competition	0.8887	-2.50**	Ambigous	
FirstEsteemed	0.9952	-0.10	Dec	
Teamsize	0.9937	-1.52		
Age	0.9855	-0.26		
AgeSq	1.0000	0.00		
Professor	0.9761	-0.25		
Responsible	0.9953	-3.73***		Dec
OpenExchange	1.3048	5.30***		Inc
Publications	1.0013	1.30		
Basic	1.1990	3.92***		Inc
OwnResearch	1.0018	0.67		
Patents	0.9468	-2.30**		Dec
Consult	0.9770	-2.60***		Dec
FamilyEnt	0.8251	-1.98**		Dec
Married	1.0900	0.84		Dec
Male	0.7389	-3.06***		
UK	0.8777	-1.25		
ExploitLose	0.9270	-1.85*	Inc	
Question & field fixed effects				
r-Square	0.0706			
Obs.	3063			

***** Significant at 1% ** Significant at 5% * Significant at 10%**

Willingness to Share: General *vs.* Specific



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)