

Network
structure and
outcomes

Paul Schrimpf

Introduction

Describing
Networks

Network
structure &
outcomes

References

Network structure and outcomes

Paul Schrimpf

UBC
Vancouver School of Economics

April 1, 2025

Introduction

Describing
Networks

Network
structure &
outcomes

References

References

- Overviews:
 - Jackson (2010),
<https://class.coursera.org/networksonline-001>
 - Goyal (2012)
- Network industries: Economides (1996), Economides and Encaoua (1996)

Introduction

Describing
Networks

Network
structure &
outcomes

References

1 Introduction

2 Describing Networks

3 Network structure & outcomes

Fershtman and Gandal (2011)

Claussen, Falck, and Grohsjean (2012)

Related Subsequent Work

Zacchia (2019)

Acemoglu, Akcigit, and Kerr (2016)

Network
structure and
outcomes

Paul Schrimpf

Introduction

Describing
Networks

Network
structure &
outcomes

References

Section 1

Introduction

- Network: nodes & links between them
- Questions:
 - How does network structure affect behavior?
 - How are networks formed?
- Focus on networks that are not owned by a single entity
 - i.e. not on network industries where a single firm owns and controls its network (telecom, electricity, airlines, etc)
 - Relatively new area, little empirical work

Introduction

Describing
Networks

Network
structure &
outcomes

References

- R&D collaboration
- Trade
- Buyer-supplier
- Consumer information & targeting

Network
structure and
outcomes

Paul Schrimpf

Introduction

Describing
Networks

Network
structure &
outcomes

References

Section 2

Describing Networks

Introduction

Describing
Networks

Network
structure &
outcomes

References

- Nodes $\in \{1, \dots, N\} = \mathcal{N}$
- Adjacency matrix G , $N \times N$ matrix with g_{ij} representing connection between i and j
- Graph $\equiv (\mathcal{N}, G)$
 - Undirected \equiv symmetric G
 - Directed \equiv asymmetric G
 - Unweighted (or discrete) $\equiv g_{ij} \in \{0, 1\}$
 - Weighted \equiv not discrete

Summary statistics 1

- Distance between two nodes = shortest path between them (∞ if no connected path)
- Diameter = largest distance between nodes
- Clustering
 - of graph is portion of j and k connected given j and k both connected to i
 - of node i is the portion of time j and k are connected directly given j and k are connected to i
 - average clustering of graph = average across nodes of node clustering
- Degree of a node = number of links
 - Directed graphs: in-degree & out-degree
 - Degree centrality =
$$\frac{\text{degree}}{N-1}$$
 - Network density = fraction of possible links present
$$= \frac{\text{average degree}}{N-1}$$
 - Degree distribution = CDF of node degree
- Centrality measures:

Introduction

Describing
Networks

Network
structure &
outcomes

References

Summary statistics 2

- Degree, closeness, betweenness, decay
- Eigenvector, Katz, Bonacich

Network
structure and
outcomes

Paul Schrimpf

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

Section 3

Network structure & outcomes

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

Network structure & outcomes

- Question: how does network structure affect some outcome?
- Reduced form work: regress outcome on node or network summary statistics

Example: Fershtman and Gandal (2011)

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

- Knowledge spillovers in open-source projects
- Data:
 - Sourceforge
 - Contributor network: linked if participated in same project
 - Project network: linked if have common contributors
- Question: how important are project vs contributor spillovers for project success?
 - Project spillover = developers learn from working on a particular project
 - Contributor spillover = developers learn from working with other developers
- Related paper: [Claussen, Falck, and Grohsjean \(2012\)](#)

Example: Fershtman and Gandal (2011)

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

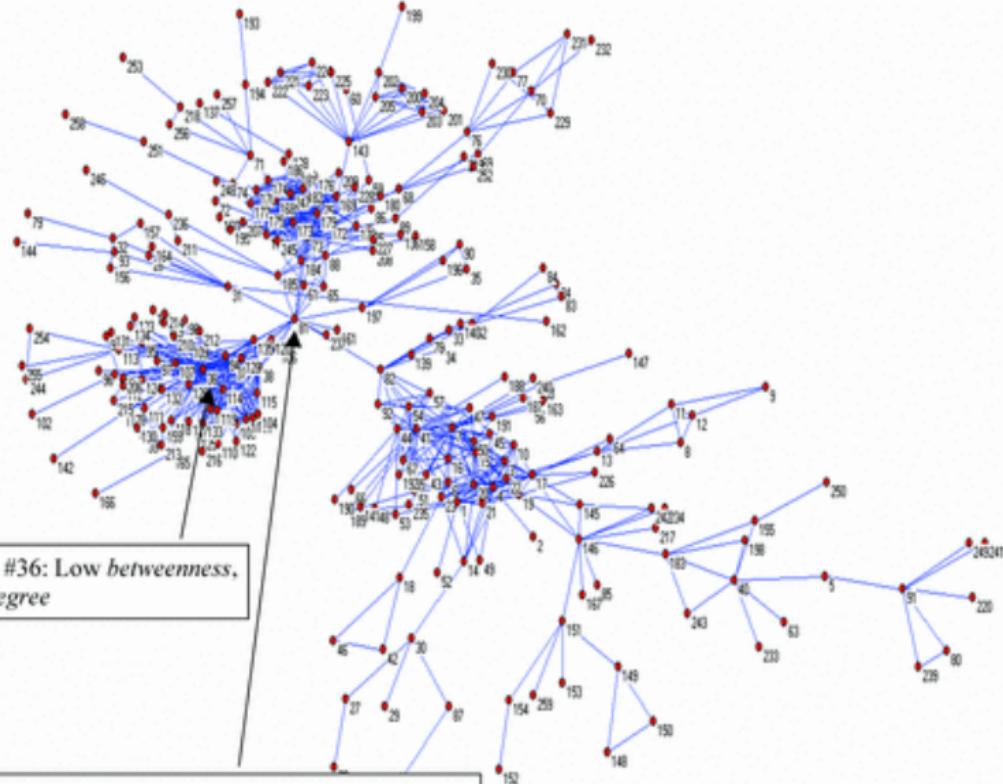
Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References



Project #36: Low betweenness,
High degree

Project #81: High betweenness, relatively low degree

Example: Fershtman and Gandal (2011)

Introduction

Describing
NetworksNetwork
structure &
outcomesFershtman and Gandal
(2011)Claussen, Falck, and
Grohsjean (2012)Related Subsequent
Work
Zacchia (2019)Acemoglu, Akcigit, and
Kerr (2016)

References

Table: The Distribution of Contributors per Project and Projects per Contributor

Project Network		Contributor Network	
Contributors	N Projects	Projects	N Contributors
1	77,571	1	123,562
2	17,576	2	22,690
3–4	11,362	3–4	10,347
5–9	6,136	5–9	3,161
10–19	1,638	10–19	317
20–49	412	20–49	26
≥ 50	56	≥50	1
Total projects	114,751	Total contributors	160,104

Example: Fershtman and Gandal (2011)

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

Table: Distribution of Component Size

Component Size (Contributors)	Components (Subnetworks)
55,087	1
196	1
65–128	2
33–64	27
17–32	152
9–16	657
5–8	2,092
3–4	4,810
2	8,287
1	47,787

Example: Fershtman and Gandal (2011)

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

Table: Distribution of Degree

Degree	Number of Contributors
0	47,787
1	22,133
2	14,818
3–4	20,271
5–8	20,121
9–16	16,228
17–32	10,004
33–64	5,409
65–128	2,040
129–256	802
257–505	491

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

- Degree centrality as measure of direct connections
- Closeness centrality = $C_c(i) = \frac{N-1}{\sum_{j \in N} d(i,j)}$ – conditional on degree measures indirect connections
- $S_i = \text{success} = \text{number of downloads}$

$$S_i = \alpha + \gamma C_c(i)/(N - 1) + \beta \text{degree}_i + \text{controls}$$

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Ace moglu, Akcigit, and
Kerr (2016)

References

Results - project network

Table: Regression Results: Dependent Variable: ldownloads

Independent variables	Regression 1		(All 66,511 Projects)		Regression 2	(Giant Component: 18,697 Projects)	
	Coefficient	T-Statistic	Coefficient	T-Statistic		Coefficient	T-Statistic
Constant	0.72	17.76	1.45	3.62			
lyears_since	1.42	60.66	1.68	31.08			
lcount_topics	0.23	9.07	0.18	3.59			
lcount_trans	0.35	11.73	0.45	8.15			
lcount_aud	0.36	10.44	0.44	5.85			
lcount_op_sy	0.11	5.95	0.18	5.00			
ds_1	-1.96	-60.57	-2.01	-31.90			
ds_2	-0.60	-17.58	-0.78	-11.50			
ds_3	0.89	25.83	0.66	9.95			
ds_4	1.86	57.21	1.80	29.27			
ds_5	2.72	79.97	2.61	40.96			
ds_6	2.12	27.07	2.03	15.35			
inactive	0.45	6.11	0.39	2.75			
lcpp	0.46	18.71	0.87	29.34			
ldegree	0.19	9.45	0.079	2.10			
giant_comp	-0.21	-3.86					
lgiant_cpp	0.44	12. 05					
lgiant_degree	-0.05	-1.26					
lcloseness			0.69				3.21
Number of observations	66,511		18,697				
Adjusted R ²	0.41		0.40				

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

- Regress downloads on average contributor degree and average contributor closeness centrality (and controls)
- Result:
 - Coefficient on log average closeness = 0.12, with $t = 1.59$
 - Coefficient on log average degree = -0.019 , with $t = -0.72$
- Including both contributor and project measures, project ones significant, contributor ones not

Network structure and outcomes

Robustness

Paul Schrimpf

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

Dept Variable: Ldownloads	$\geq 2\text{yr}$, Coef	$\geq 200\text{dl}$ T-Stat	$< 3.6\text{yr}$, Coef	$\geq 200\text{dl}$ T-Stat	$\geq 2\text{yr}$, Coef	$\geq 200\text{dl}$ T-Stat
Independent variables						
Constant	5.75	13.35	6.21	14.55	8.51	16.29
lyears_since	1.08	11.18	0.91	11.40	1.06	10.97
lcount_topics	-0.06	-1.31	-0.06	-1.12	-0.06	-1.23
lcount_trans	0.42	9.61	0.44	8.83	0.41	9.39
lcount_aud	0.21	2.65	0.46	5.62	0.18	2.28
lcount_op_sy	0.26	7.91	0.26	6.62	0.26	7.94
ds_1	-0.46	-6.83	-0.75	-6.23	-0.46	-6.88
ds_2	-0.57	-7.31	-0.52	-4.80	-0.57	-7.68
ds_3	-0.27	-4.35	-0.23	-2.72	-0.28	-4.44
ds_4	0.19	3.45	0.18	2.41	0.18	3.24
ds_5	0.75	12.99	0.54	6.92	0.73	12.80
ds_6	0.73	7.00	0.45	3.06	0.72	6.89
Inactive	-0.018	-0.12	0.02	0.11	-0.041	-0.28
lcpp	0.76	22.21	0.63	19.30	0.59	15.38
ldegree	0.19	5.07	0.0038	0.09	-0.019	-0.43
lcloseness	0.71	3.28	0.54	2.37	0.45	2.08
lbetweenness					0.30	9.21
Number of observations	6,397		4,086		6,397	
Adjusted R ²	0.28		0.25		0.29	

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

- How to interpret results?
 - Network structure affects downloads
 - Downloads affect contributions, which affects network structure
- Why closeness centrality and degree? (they do explore robustness to other measures, but none of them theoretically motivated)

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

Claussen, Falck, and Grohsjean (2012)

- Developer networks in electronic games
- Panel data 1972-2007 on games & developers
- Construct network of developers 1995-2007
- Developers linked in year t if worked together anytime between 1972 and t
- Look at relationship between revenue (or rating) & degree centrality & closeness centrality

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

Claussen, Falck, and Grohsjean (2012)

$$S_{igdpt} = \alpha_i + \alpha_d + \alpha_p + \alpha_t + \beta_1 D_{igdpt-1} + \beta_2 C_{igdpt-1} + CV_{igdpt}\gamma + \epsilon_{igdpt}$$

- Developer i
- Game g
- Developing firm d
- Publisher p
- Year t

Introduction

Describing
NetworksNetwork
structure &
outcomesFershtman and Gandal
(2011)Claussen, Falck, and
Grohsjean (2012)Related Subsequent
Work
Zacchia (2019)Acemoglu, Akcigit, and
Kerr (2016)

References

Table 1
Summary statistics.

Variable	N	Mean	SD	Min	Max
In(revenue)	151,677	14.958	1.701	4.264	19.440
Critics' score	146,675	0.007	0.781	-3.831	2.223
Degree centrality D_{igdpt}	148,627	0.001	0.002	0.000	0.041
Closeness centrality C_{igdpt}	148,627	0.205	0.038	0.052	0.338
Leading position	151,677	0.213	0.410	0	1
Tenure	151,677	3.871	4.254	0	28
Team size	151,677	65.780	53.234	1	297
Licensed game	151,677	0.362	0.480	0	1

Claussen, Falck, and Grohsjean (2012)

Table 3

Baseline regression results with revenue as success measure.

	(3-1)	(3-2)	(3-3)	(3-4)
	Dependent variable: ln(revenue)			
Degree centrality D_{igdpt}	8.494* (4.468)	8.029* (4.342)	7.281** (3.321)	6.512* (3.644)
Closeness centrality C_{igdpt}	-0.137 (0.201)	-0.307 (0.217)	-0.105 (0.155)	-0.223 (0.174)
Co-worker degree c. \bar{D}_{-igdpt}			40.20 (60.77)	56.53 (45.38)
Co-worker closeness \bar{C}_{-igdpt}			-0.548 (3.059)	-3.122 (2.567)
Tenure	0.0170 (0.0618)	0.0612 (0.0886)	0.0236 (0.0664)	0.0676 (0.0932)
Team size	0.00447*** (0.000937)	0.00396*** (0.000990)	0.00446*** (0.000938)	0.00394*** (0.000991)
Licensed game	0.192*** (0.0720)	0.172** (0.0771)	0.193*** (0.0721)	0.171** (0.0774)
Network measures lagged	No	Yes	No	Yes
Observations	151,484	94,597	151,443	94,388
Number developers	56,944	30,993	56,937	30,956
Within-developer R^2	0.635	0.638	0.635	0.638
Between-developer R^2	0.802	0.742	0.798	0.736
Overall R^2	0.736	0.689	0.734	0.684

Notes: Fixed-effect OLS point estimates with standard errors clustered at the project level in parentheses. As the panels are not nested within the clusters, a degree of freedom adjustment is conducted, producing conservative results for the standard errors. Asterisks denote significance levels (** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$). All specifications control for fixed effects on the level of the developer, the developing

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

“Network dynamics and knowledge transfer in virtual organisations”

- Panel data of Sourceforge contributions
- Look for direct & indirect spillovers
- Programmers who work on many projects positively impact success beyond their effect on connectivity in the network

Network
structure and
outcomes

Paul Schrimpf

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

Delcamp and Leiponen (2014)

"Innovating standards through informal consortia: The case of wireless telecommunications"

Network
structure and
outcomes

Paul Schrimpf

Branstetter, Gandal, and Kuniesky (2017)

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

“Network-Mediated Knowledge Spillovers: A Cross-Country Comparative Analysis of Information Security Innovations”

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

“Dynamics of Open Source Movements”

- Dynamic model of open source contributions and commercial competitors
- Theory paper, not empirical

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

Knowledge Spillovers through Networks of Scientists

Zacchia (2019)

- Weighted network of publicly traded companies
- Links = proportion of firms' inventors that have former patent collaborations
- Main endogeneity concern: common unobservables
- IV motivated by model of firm interaction

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

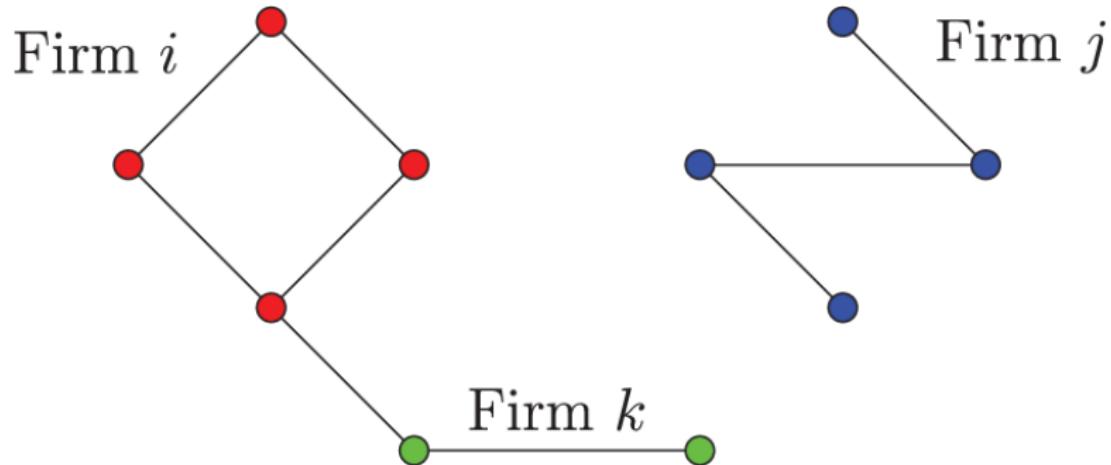
Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

Network Among Inventors

- Inventor m and n linked ($p_{(mn)t} = 1$) if m and n collaborated on any past patent



Network Among Firms

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

$$c_{(ij)t}^f = f \left(\frac{\text{inv. of } i \text{ connected to } j \text{ at } t + \text{inv. of } j \text{ connected to } i \text{ at } t}{\text{inv. of } i \text{ at } t + \text{inv. of } j \text{ at } t} \right)$$

- Symmetric
- $0 \leq c_{(ij)t}^f \leq 1$
- In empirical results, $f(\cdot) = \sqrt{\cdot}$ and $g_{(ij)t} = c_{(ij)t}^{\sqrt{\cdot}}$

Introduction

Describing Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

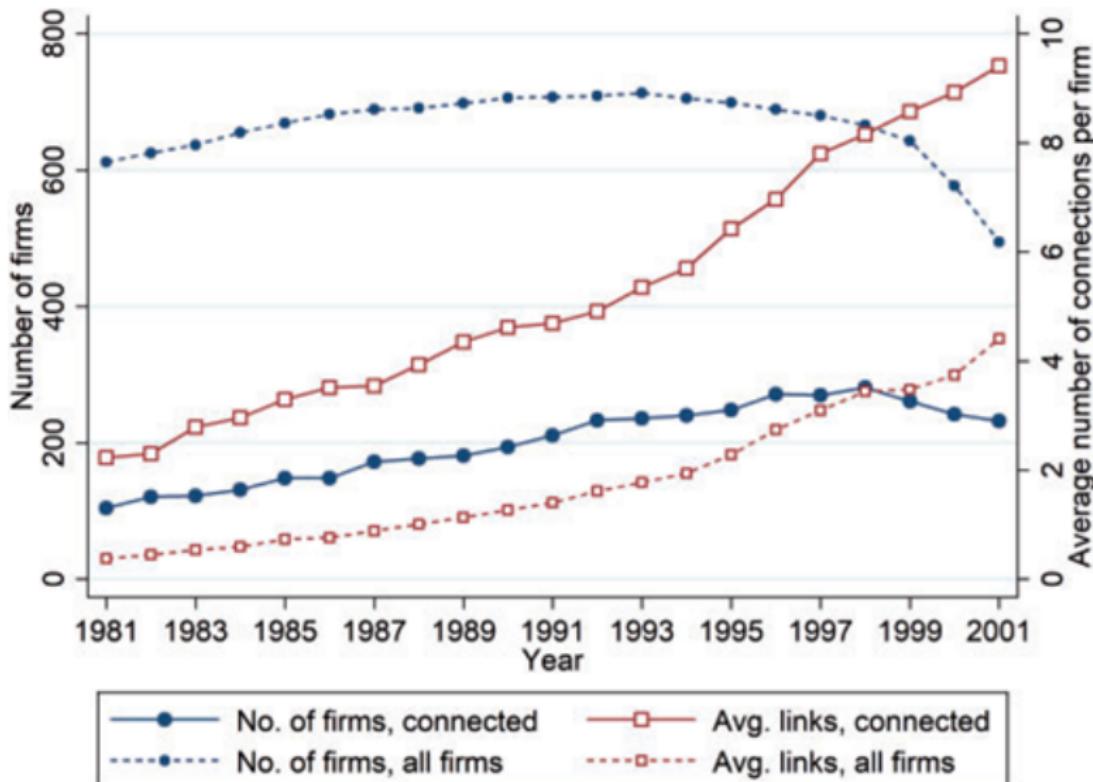
Claussen, Falck, and
Grohsjean (2012)

Related Subsequent Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References



Introduction

Describing Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

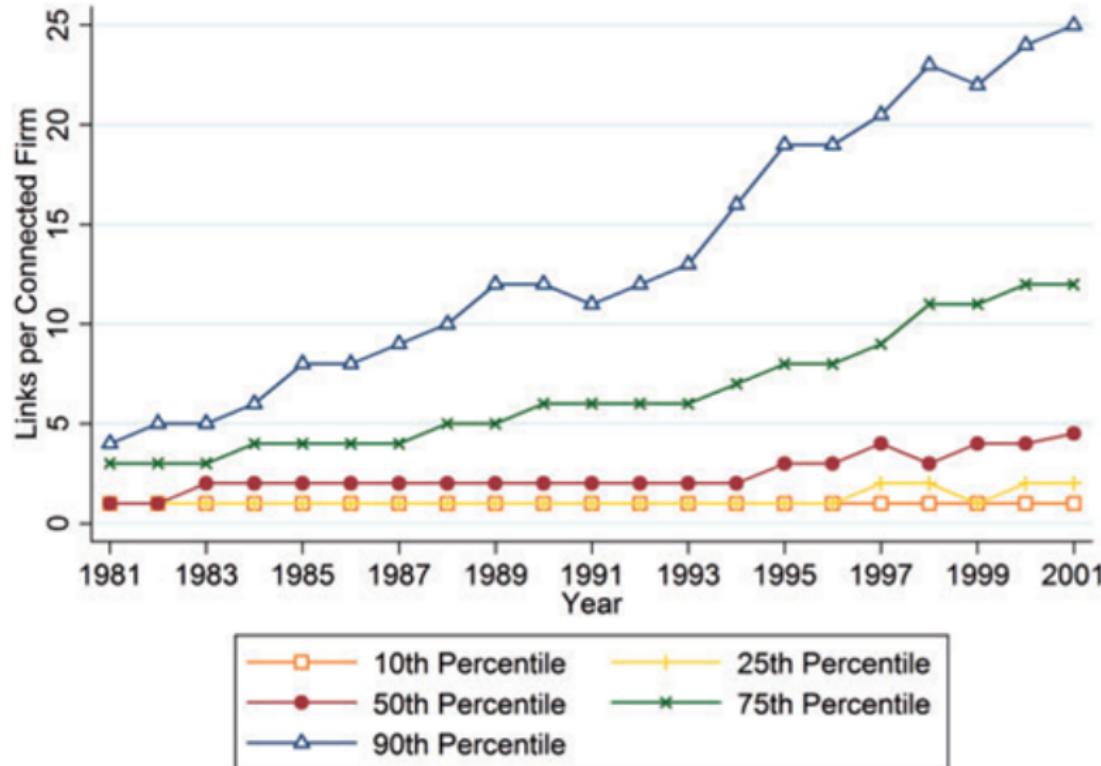
Claussen, Falck, and
Grohsjean (2012)

Related Subsequent Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References



Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

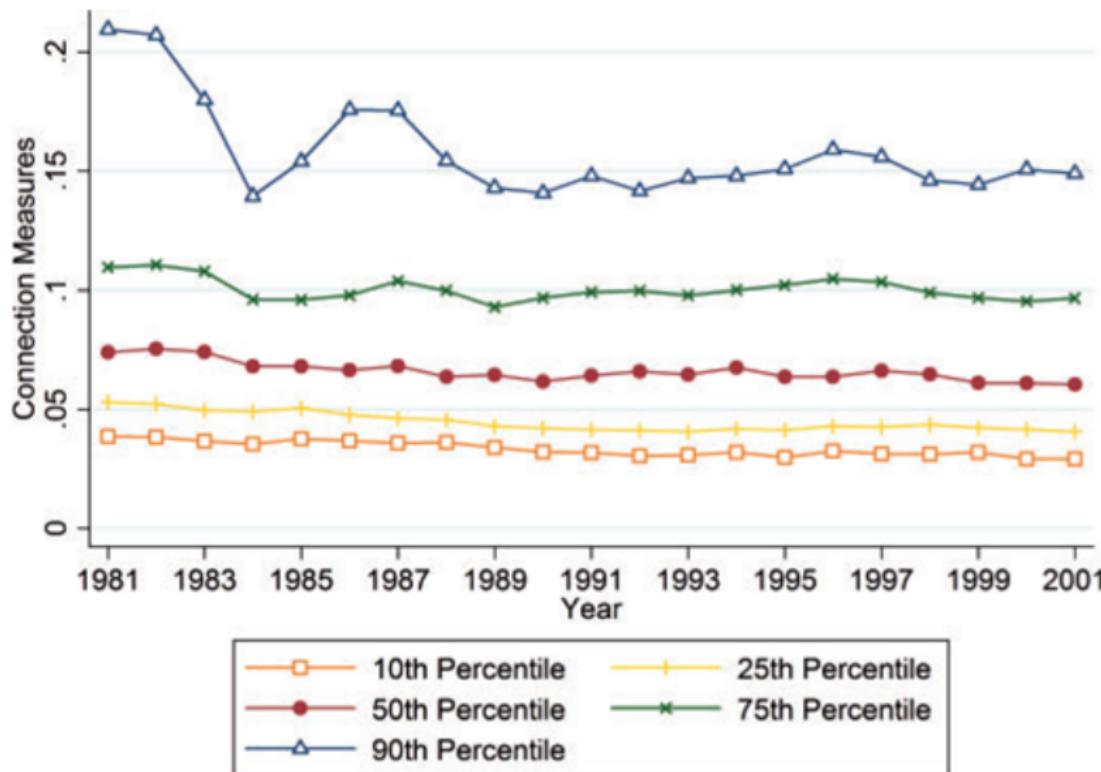
Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References



Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

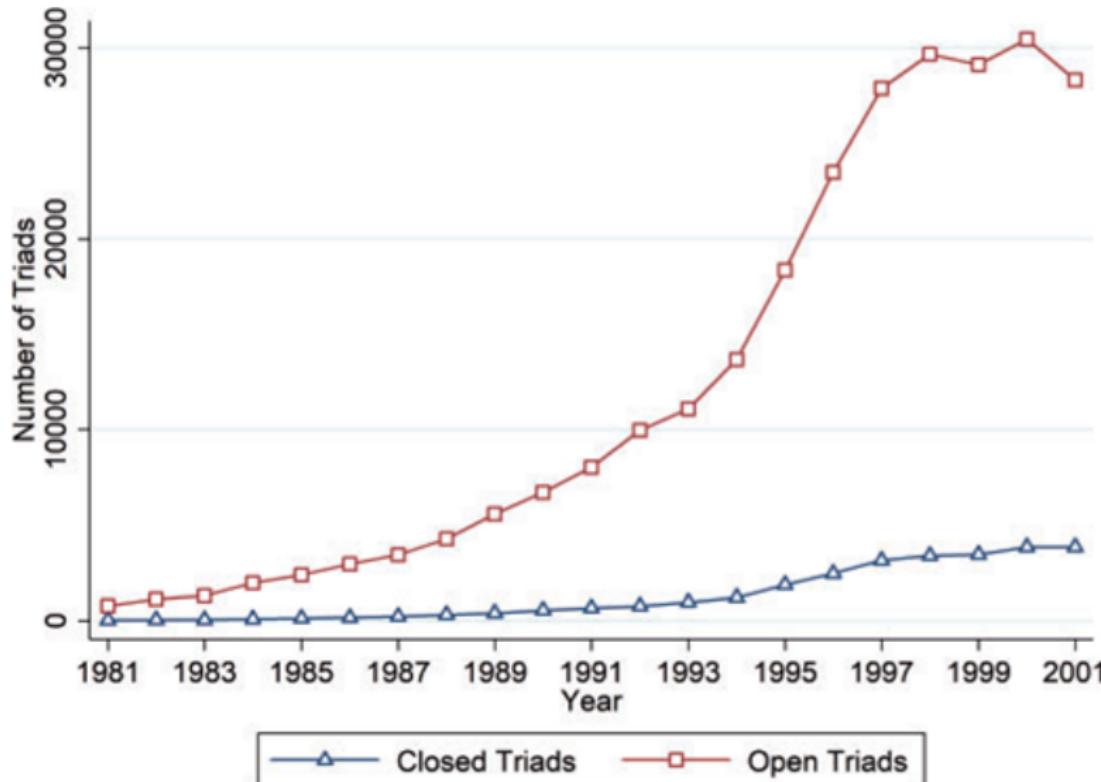
Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References



Network structure and outcomes	TABLE 1					
	Summary statistics, 1981–2001					
Paul Schrimpf		No network	Quartile of $\sum_t \bar{g}_{it}$			
			1	2	3	
Introduction	Y_{it} : real sales (Millions 1996\$\\$)	751	1,066	1,383	2,172	10,462
		(3,792)	(2,357)	(2,504)	(4,533)	(20,058)
	V_{it}/A_{it} : Tobin's q	1.886	1.885	2.573	2.734	3.410
		(2.031)	(1.839)	(3.080)	(3.306)	(4.118)
	P_{it} : patent stock (cit. weighted)	7.453	16.09	24.65	74.03	652.0
		(48.17)	(44.75)	(50.91)	(143.8)	(1322.1)
	E_{it} : employees (thousands)	4.068	6.940	9.328	12.40	57.09
		(12.52)	(15.80)	(16.63)	(22.43)	(96.80)
	Y_{it}/E_{it} : labour productivity	135.6	134.5	157.1	156.5	192.4
		(80.06)	(106.6)	(95.43)	(117.7)	(153.3)
Describing Networks	Y_{it}/K_{it} : capital productivity	6.932	5.308	5.142	4.941	4.184
		(6.083)	(3.167)	(3.992)	(3.292)	(2.883)
Network structure & outcomes	Y_{it}/S_{it} : productivity of R&D	39.31	19.71	51.10	11.12	4.342
		(134.1)	(70.47)	(479.9)	(34.46)	(3.932)
Fershtman and Gandal (2011)	Y_{it} / Jaffe measure (i, t)	80.28	107.7	140.0	211.6	962.5
		(407.7)	(238.5)	(264.9)	(435.4)	(1787.8)
Claussen, Falck, and Grohsjean (2012)	$Y_{it}/\prod_j S_{jt}^{g_{j it}}$: Y to spillover pool	953.9	846.2	577.6	198.9	
		(2,224.0)	(1,762.1)	(1,858.7)	(1,339.6)	
Related Subsequent Work	No. of observations	4,363	1,854	1,819	1,949	2,028
Zacchia (2019)						
Acemoglu, Akcigit, and Kerr (2016)						
References						

Network structure and outcomes

Paul Schrimpf

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

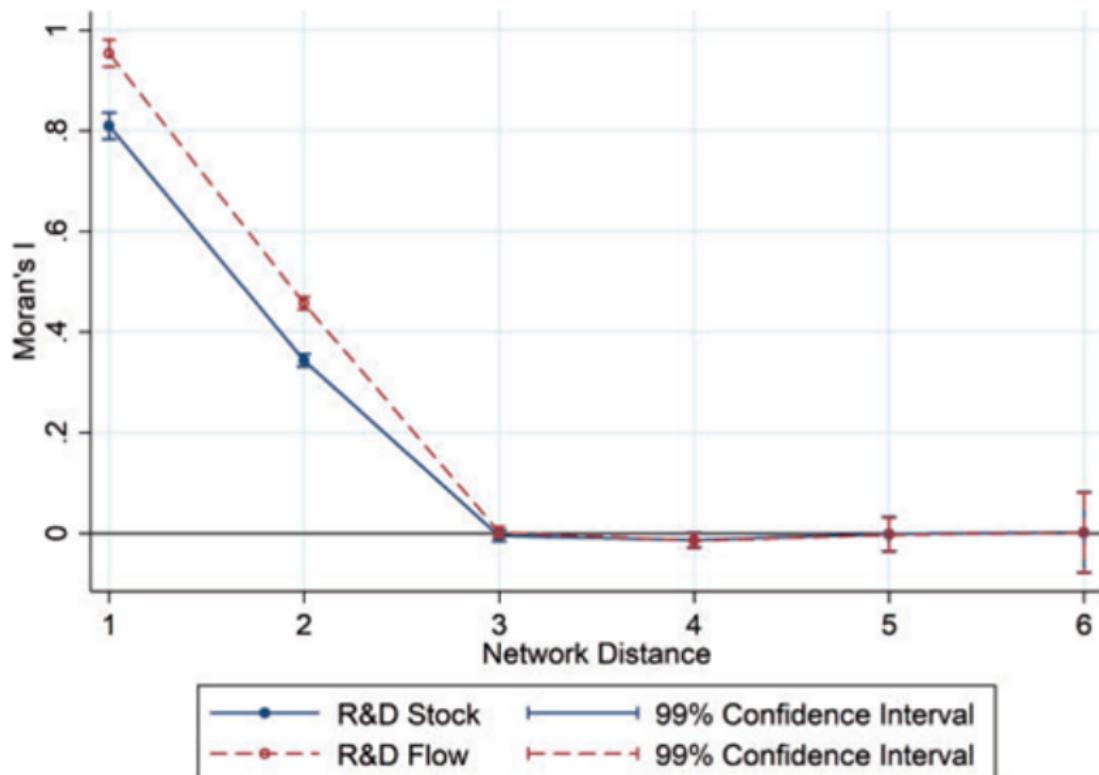
Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References



Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

$$\log Y_{it} = \alpha_i + \sum_{q=1}^Q \beta_q \log X_{itq} + \gamma \log S_{it} + \delta \sum_{j=1}^N g_{(ij)t} \log S_{jt} + \tau_t + \nu_{it}$$

- Output Y_{it}
- Inputs X_{itq}
- R&D stock S_{it} (depreciated past sum of R&D expenditures)
- δ = strength of R&D spillovers

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

- $E[\log S_{jt} v_{it}] \neq 0$ from e.g. v_{jt} correlated with v_{it} , or S_{jt} chosen with some knowledge of v_{it}
- Endogenous connections $E[g_{(ij)t} v_{it}] \neq 0$

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

- Firms \mathcal{I} with connection \mathcal{G}
- Knowledge capital

$$\tilde{s}_i = s_i^\gamma \left(\prod_j s_j^{g_{ij}} \right)^\delta$$

- R&D cost $e^{\bar{\omega}_i} s_i$
- Cobb-Douglas Production as above
- Firms maximize profits (output minus linear input costs minus R&D costs)

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

- Unique Bayes-Nash equilibrium with

$$\log s_i^* = \frac{\log \gamma + \sum_q \beta_q (\log \beta_q - \log \xi_q - \log \gamma)}{1 - \gamma - \sum_q \beta_q} b_i^*(\mathcal{G}; \vartheta) + s_i^*(\Omega_i; \mathcal{G})$$

where

- $\vartheta = \frac{\delta}{1 - \gamma - \sum_q \beta_q}$
- $b_i^*(\mathcal{G}; \vartheta)$ is Katz-Bonacich network centrality
- Ω_i is firm's information set and

$$s_i^*(\Omega_i; \mathcal{G}) = \frac{\omega_i - (1 - \sum_q \beta_q) \bar{\omega}_i + \log E[\prod_j e^{g_{ij} \delta s_j^*(\Omega_j, \mathcal{G})} | \Omega_i]}{1 - \gamma - \sum_q \beta_q}$$

Identifying Assumptions

- Correlation of $\log S_{jt}$ with unobservables of i happens through $s_i^*(\Omega_i; \mathcal{G})$ due to ω_{it} and ω_{jt} possibly being correlated and Ω_i potentially being informative about ω_{jt}
- Assumption 1: $\exists C > 0$ such that if distance from i to j is greater than C , then $\text{Cov}(\omega_i, \omega_j | d_{ij} > C) = 0$ and $\text{Cov}(\bar{\omega}_i, \bar{\omega}_j | d_{ij} > C) = 0$
- Assumption 2: $\exists L > 0$ such that $d_{ij} > L$ implies $(\omega_j, \bar{\omega}_j) \notin \Omega_i$
- Implies:

$$\text{Cov}(\omega_i, \log S_j | d_{ij} > C + L) = 0$$

$$\text{Cov}(\log S_i \log S_j | d_{ij} \leq C + 2L) \lessgtr 0$$

$$\text{Cov}(\log S_i \log S_j | d_{ij} > C + 2L) = 0$$

- Use $\log S_k$ as instrument for $\log S_j$ when k and j are distance $C + L < D \leq C + 2L$ apart

TABLE 2
Production function, ordinary least squares estimates, 1981–2001

		(1)	(2)	(3)	(4)	(5)
Introduction	Private R&D (γ)	0.0455*** (0.0108)	0.0438*** (0.0105)	0.0568*** (0.0118)	0.0554*** (0.0128)	0.0515*** (0.0142)
Describing Networks	R&D spillovers (δ)	0.0159*** (0.0023)	0.0147*** (0.0024)	0.0114*** (0.0024)	0.0118*** (0.0025)	0.0116*** (0.0028)
Network structure & outcomes	Geographic spillovers	0.0035 (0.0021)	0.0027 (0.0020)	0.0023 (0.0023)	0.0015 (0.0019)	
Fershtman and Gandal (2011)	Capital	0.2071*** (0.0143)	0.2061*** (0.0145)	0.2035*** (0.0194)	0.2020*** (0.0213)	0.2023*** (0.0198)
Claussen, Falck, and Grohsjean (2012)	Labour	0.6550*** (0.0241)	0.6580*** (0.0249)	0.6634*** (0.0351)	0.6613*** (0.0359)	0.6622*** (0.0363)
Related Subsequent Work	Jaffe tech. proximity	0.1352** (0.0581)	0.0361 (0.0583)	0.0026 (0.0766)	0.0179 (0.0805)	
Zacchia (2019)	Fixed effects	YES	YES	YES	YES	YES
Acemoglu, Akcigit, and Kerr (2016)	Only network	NO	NO	YES	YES	YES
References	No. of communities (Community \times Year Effects)	0	0	0	10	20
	No. of observations	12,503	12,503	7,607	7,607	7,607

Notes: The table reports OLS estimates of model (4.11). Columns 1 and 2 are estimated over the entire original sample of 736 firms in the time interval 1981–2001. Estimates in columns 3, 4, and 5 restrict the sample to firms with at least one non-zero connection ($g_{(ij)t} \neq 0$) in any year t ; all observations of these firms are also included for years with no connections. All estimates include firm and year fixed effects. Columns 4 and 5 include additional community-by-year fixed effects, where communities are obtained via the Louvain algorithm with $\varphi=0.8$ (10 communities) in column 4 and $\varphi=0.6$ (20 communities) in column 5. Standard errors are clustered by the 20 “communities” obtained via the Louvain algorithm with $\varphi=0.6$ (small sample corrections are applied). All observations of the same individual firm in different years enter the same cluster. For estimates not restricted to the network, firms outside the network constitute single clusters. Asterisks denote conventional significance levels of t -tests (** $p < 0.05$, *** $p < 0.01$).

TABLE 3
Production function, first stage estimates, 1981–2001

	(1)	(2)	(3)	(4)	(5)
Introduction	Distance 2 instrument 0.0043*** (0.0002)	0.0044*** (0.0002)			
Describing Networks	Distance 3 instrument 0.0001* (0.0001)	0.0006*** (0.0001)	0.0005*** (0.0001)	0.0005*** (0.0001)	0.0005*** (0.0001)
Network structure & outcomes	Private R&D 0.1598*** (0.0517)	0.1774*** (0.0488)	0.5022*** (0.1463)	0.4612*** (0.1436)	0.4357*** (0.1504)
Fershtman and Gandal (2011)	Capital 0.1827 (0.1133)	0.1895 (0.1098)	0.6169** (0.2622)	0.5903** (0.2719)	0.5923** (0.2746)
Claussen, Falck, and Grohsjean (2012)	Labour −0.0761 (0.0921)	−0.0530 (0.0913)	−0.6482** (0.2412)	−0.6021** (0.2636)	−0.6541** (0.2631)
Related Subsequent Work	Jaffe tech. proximity 1.8439*** (0.4808)	1.8224*** (0.5123)	3.3663** (1.5953)	3.2740** (1.3777)	3.3672** (1.4690)
Zacchia (2019)	Fixed effects YES	YES	YES	YES	YES
Acemoglu, Angrist, and Kerr (2016)	Only network YES	YES	YES	YES	YES
References	No. of communities (Community × Year Effects) 0	0	0	10	20
	<i>F</i> -statistic 255.17	219.47	24.92	32.06	19.18
	No. of observations 7,607	7,607	7,607	7,607	7,607

Notes: The table reports OLS “first stage” regressions of the spillover variable $\sum_{j \neq i} g_{(ij)t} \log S_{jt}$ on selected instruments and all other right-hand side variables included in the regressions from Table 2. The sample is restricted to firms with at least one non-zero connection ($g_{(ij)t} \neq 0$) in any year t ; all observations of these firms are also included for years with no connections. Columns 1 and 2 include, on the right hand side, the distance 2 instrument; columns 2 through 5 include the distance 3 instrument. All estimates include firm and year fixed effects. Columns 4 and 5 include additional community-by-year fixed effects, where communities are obtained via the Louvain algorithm with $\varphi = 0.8$ (10 communities) in column 4 and $\varphi = 0.6$ (20 communities) in column 5. Standard errors are clustered by the 20 “communities” obtained via the Louvain algorithm with $\varphi = 0.6$ (small sample corrections are applied). All observations of the same individual firm in different years enter the same cluster. Asterisks denote conventional significance levels of t -tests (* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$).

TABLE 4
Production function, two stages least squares estimates, 1981–2001

	(1)	(2)	(3)	(4)	(5)
Introduction	Private R&D (γ)	0.0560*** (0.0117)	0.0562*** (0.0118)	0.0510*** (0.0111)	0.0489*** (0.0127)
Describing Networks	R&D spillovers (δ)	0.0127*** (0.0029)	0.0125*** (0.0030)	0.0204** (0.0084)	0.0230** (0.0088)
Network structure & outcomes	Geographic spillovers	0.0027 (0.0019)	0.0027 (0.0019)	0.0030 (0.0018)	0.0026 (0.0021)
Fershtman and Gandal (2011)	Capital	0.2025*** (0.0200)	0.2027*** (0.0199)	0.1969*** (0.0225)	0.1944*** (0.0244)
Claussen, Falck, and Grohsjean (2012)	Labour	0.6642*** (0.0357)	0.6640*** (0.0356)	0.6692*** (0.0373)	0.6677*** (0.0382)
Related Subsequent Work	Jaffe tech. proximity	0.0314 (0.0549)	0.0324 (0.0553)	0.0041 (0.0545)	-0.0364 (0.0669)
Zacchia (2019)	Spillovers IV(s)	D=2	D=2,3	D=3	D=3
Acemoglu, Angrist, and Kerr (2016)	Fixed effects	YES	YES	YES	YES
References	Only network	YES	YES	YES	YES
	No. of communities				
	(Community \times Year Effects)	0	0	0	10
	No. of observations	7,607	7,607	7,607	7,607
					20
					7,607

Notes: The table reports IV-2SLS estimates of model (4.11). All estimates are restricted to firms with at least one non-zero connection ($g_{(ij)t} \neq 0$) in any year t ; all observations of these firms are also included for years with no connections. Models in columns 1 and 2 employ the distance 2 instrument; models in columns 2 through 5 employ the distance 3 instrument. All estimates include firm and year fixed effects. Columns 4 and 5 include additional community-by-year fixed effects, where communities are obtained via the Louvain algorithm with $\varphi = 0.8$ (10 communities) in column 4 and $\varphi = 0.6$ (20 communities) in column 5. Standard errors are clustered by the 20 “communities” obtained via the Louvain algorithm with $\varphi = 0.6$ (small sample corrections are applied). All observations of the same individual firm in different years enter the same cluster. Asterisks denote conventional significance levels of t -tests (* $p < 0.05$; ** $p < 0.01$).

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

- $D = 2$ IV near OLS implies $D = 2$ might be too small, i.e. $C = 2$
- Marginal social return = $(\gamma + \delta \bar{g}_i) \frac{Y_i}{S_i} = 114\%$ considerably greater than marginal private return = $\gamma \frac{Y_i}{S_i} = 102\%$

- Recent papers using credible identification strategies to measure network spillovers
 - Myers and Lanahan (2022) “Estimating Spillovers from Publicly Funded R&D: Evidence from the US Department of Energy”
 - Schnitzer and Watzinger (2022) “Measuring the Spillovers of Venture Capital”
 - Arora, Belenzon, and Sheer (2021) “Knowledge Spillovers and Corporate Investment in Scientific Research”
 - Ganguli, Lin, and Reynolds (2020) “The Paper Trail of Knowledge Spillovers: Evidence from Patent Interferences”
 - Bailey et al. (2022) “Peer Effects in Product Adoption”
- Malikov and Zhao (2023) “On the Estimation of Cross-Firm Productivity Spillovers with an Application to FDI” (modification of production function framework to allow for productivity spillovers)

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

Acemoglu, Akcigit, and Kerr (2016)

“Innovation network”

- Directed network of patent citations, 1975-2004
- Results:
 - Network stable over time
 - Past innovations (patents) in connected industries predict current patents
 - Impact of innovations are localized

Network structure and outcomes

Paul Schrimpf

Introduction

Describing Networks

Network structure & outcomes

Fershtman and Gandal
(2011)

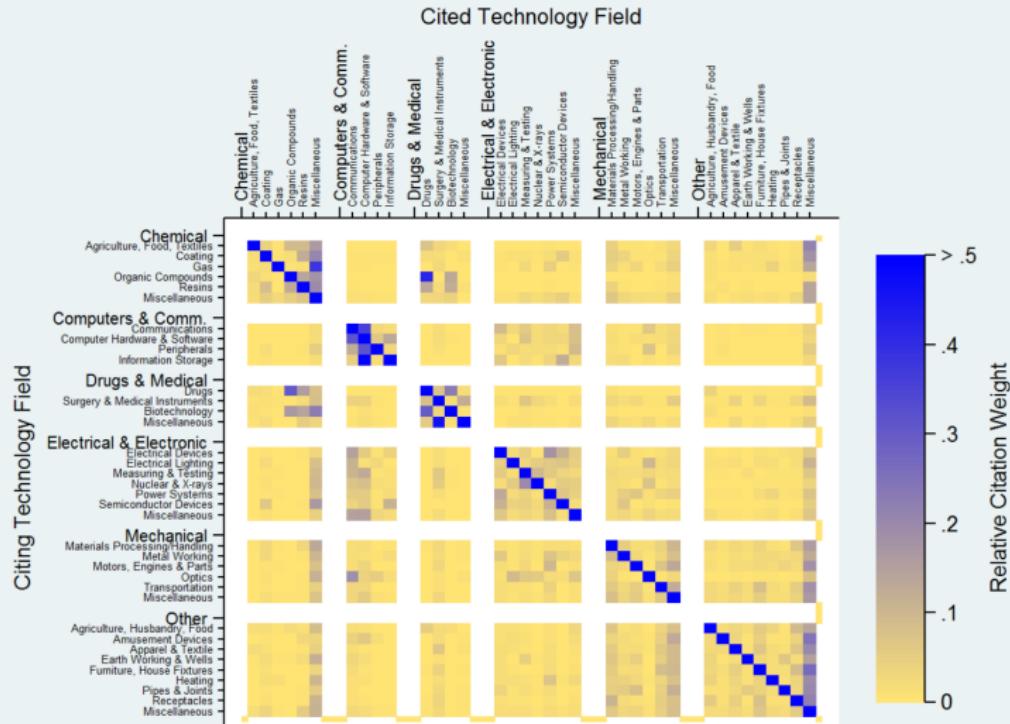
Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References



Introduction

Describing Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

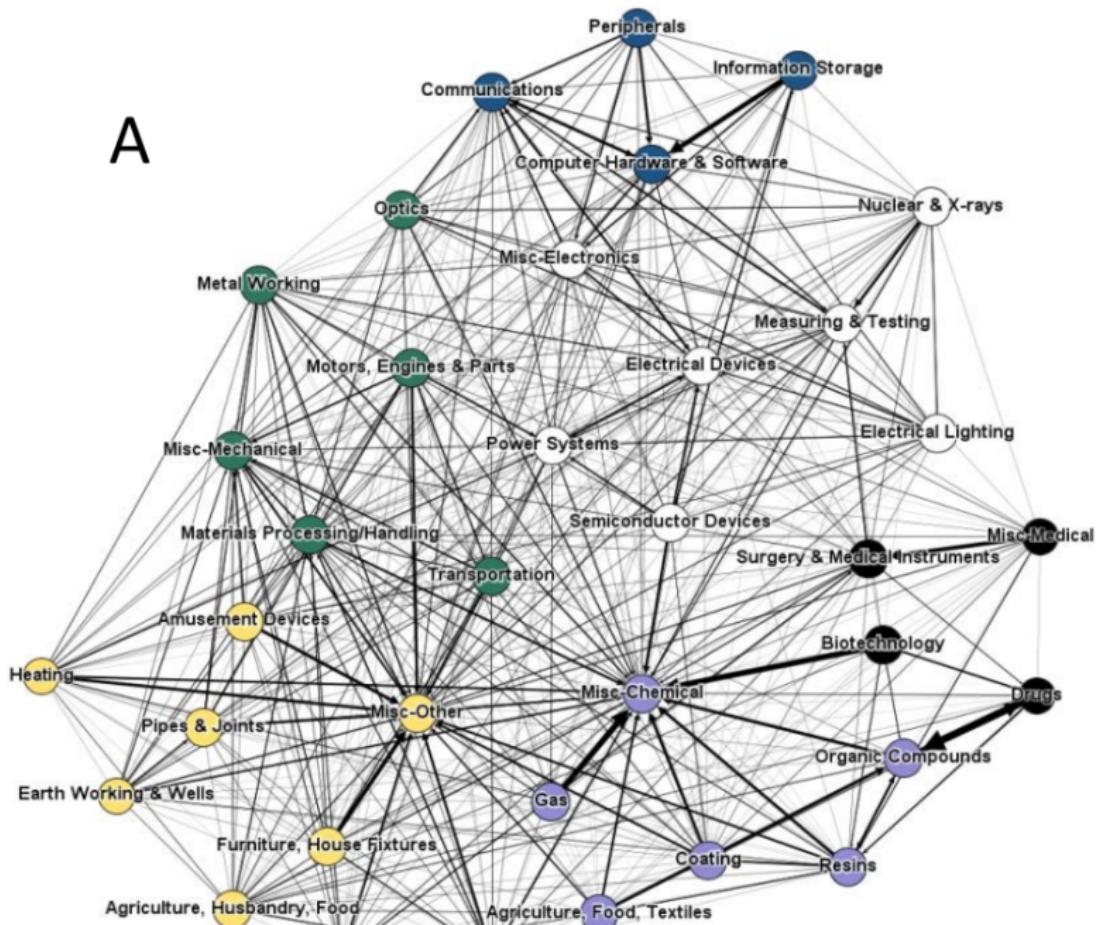
Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

A



Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

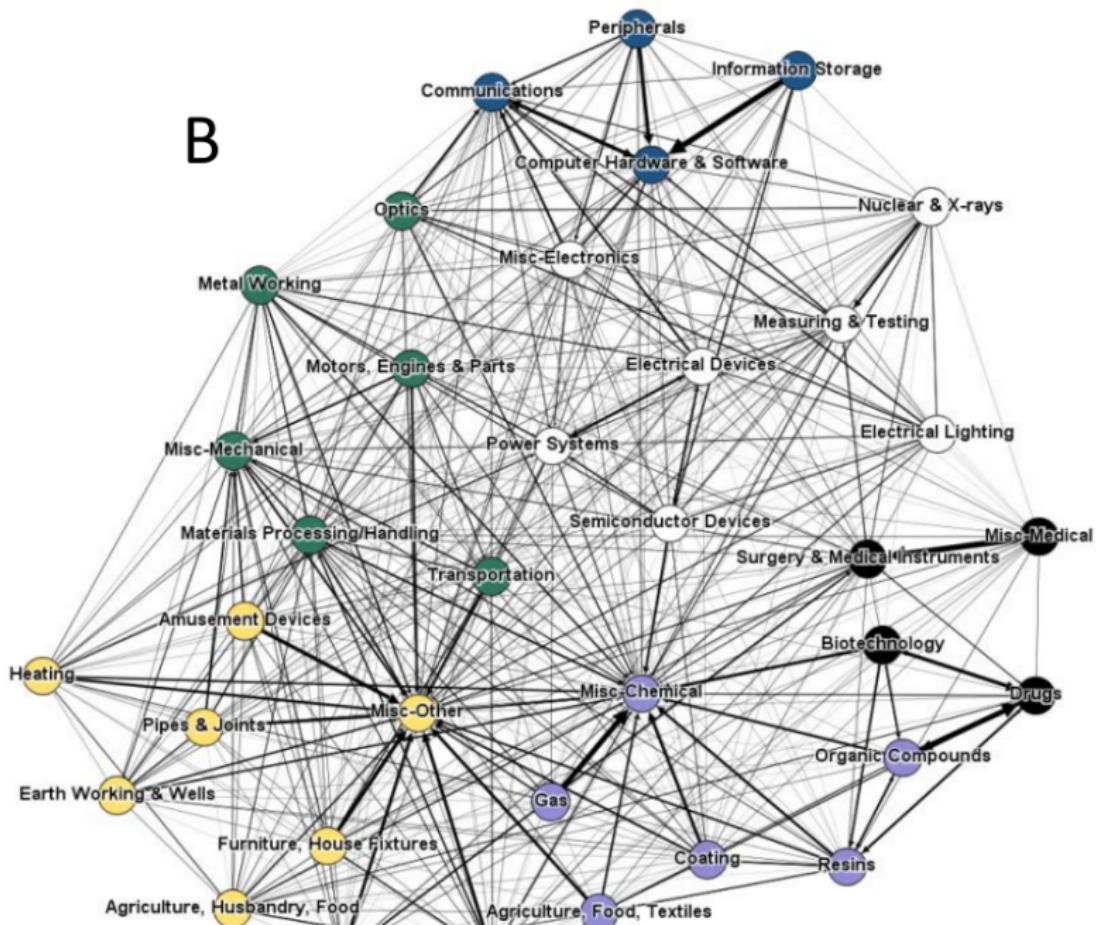
Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References



Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

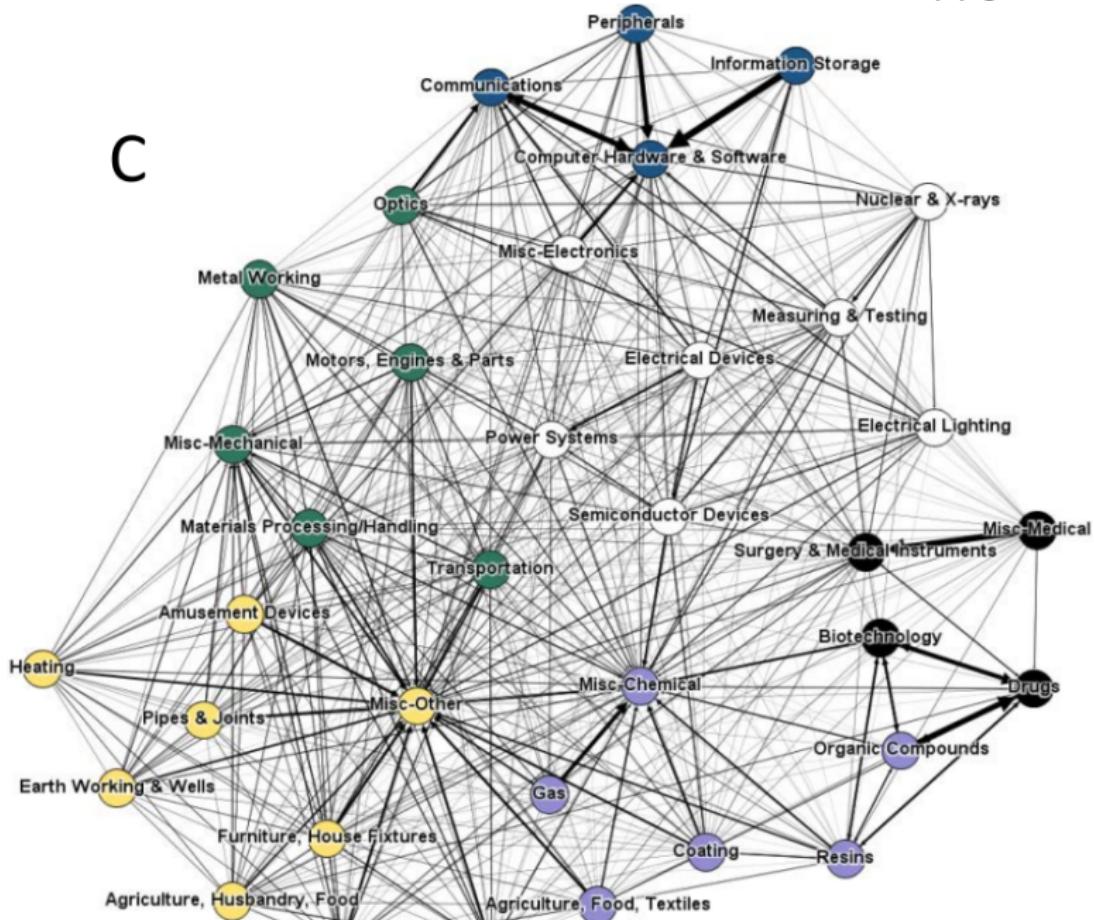
Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

1995-2004

C



Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

Predicting patents

- Predicted patents in industry j from past citations:

$$\hat{P}_{j,t} = \sum_{k \neq j} \sum_{a=1}^{10} \frac{\text{Citations}_{j \rightarrow k,a}}{\text{Patents}_k} P_{k,t-a}$$

where

- $\text{Citations}_{j \rightarrow k,a}$ = citations of a patent in industry k that is a years old from j
- Patents_k = total patents in k
- Both estimated using 1975-1994 data
- Predictions for 1995-2004

Introduction

Describing
Networks

Network
structure &
outcomes

Fershtman and Gandal
(2011)

Claussen, Falck, and
Grohsjean (2012)

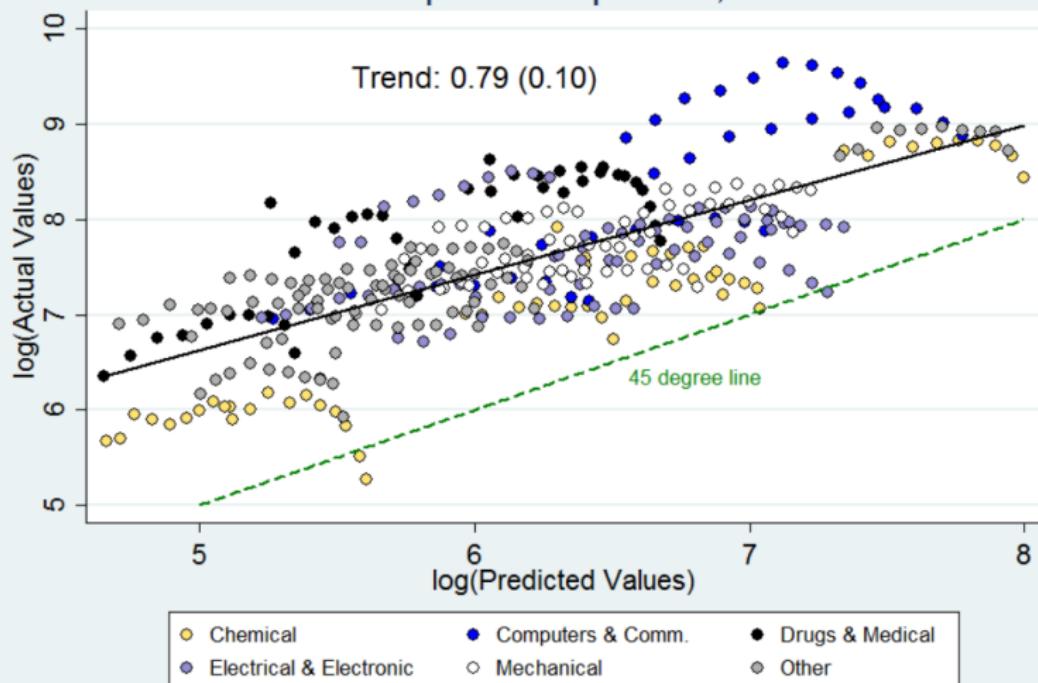
Related Subsequent
Work

Zacchia (2019)

Acemoglu, Akcigit, and
Kerr (2016)

References

A: Actual vs. predicted patents, 1995-2004



Appendix Table 2: Disaggregated analysis of innovation network

	Full sample with >5 per annum			Restricted sample with >50 per annum		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable is log cumulative patent counts in patent class 1995-2004						
Log cumulative patents 1985-1994	0.678 (0.061)	0.732 (0.055)	0.840 (0.066)	0.785 (0.083)	0.806 (0.069)	0.905 (0.079)
Log expected patenting from network stimulus	0.345 (0.052)			0.265 (0.066)		
Log expected patenting from network stimulus due to top 10 upstream classes		0.298 (0.056)			0.294 (0.066)	
Log expected patenting from network stimulus due to next 10 upstream classes		0.067 (0.074)			-0.041 (0.091)	
Log expected patenting from network stimulus outside of top 20 upstream classes		-0.061 (0.054)			-0.032 (0.068)	
Log expected patenting from network stimulus within subcategory			0.077 (0.029)			0.043 (0.037)
Log expected patenting from network stimulus within rest of category			0.105 (0.037)			0.097 (0.047)
Log expected patenting from network stimulus outside of category			-0.004 (0.048)			0.012 (0.054)

Notes: See Appendix Table 1. In Columns 2 and 5, we separate the upstream stimulus provided by the ten most-important upstream classes, the next ten upstream classes, and those beyond. These upstream classes are defined by citation shares made by the focal class during the pre-1995 period. In two cases, the patent class cites fewer than ten upstream categories and is excluded. Columns 3 and 6 alternatively rely on the USPTO classification system of subcategories and categories, which is naturally cruder since it is less-tailored to an individual technology's citation patterns. The disaggregated results do not add up to the total network effect due to the log transformations.

Network
structure and
outcomes

Paul Schrimpf

Introduction

Describing
Networks

Network
structure &
outcomes

References

- Acemoglu, Daron, Ufuk Akcigit, and William R Kerr. 2016. “Innovation network.” *Proceedings of the National Academy of Sciences* :201613559 URL <http://www.pnas.org/content/113/41/11483.abstract>.
- Arora, Ashish, Sharon Belenzon, and Lia Sheer. 2021. “Knowledge Spillovers and Corporate Investment in Scientific Research.” *American Economic Review* 111 (3):871–98. URL <https://www.aeaweb.org/articles?id=10.1257/aer.20171742>.
- Athey, Susan and Glenn Ellison. 2014. “Dynamics of Open Source Movements.” *Journal of Economics & Management Strategy* 23 (2):294–316. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/jems.12053>.
- Bailey, Michael, Drew Johnston, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. 2022. “Peer Effects in Product Adoption.” *American Economic Journal: Applied Economics* 14 (3):488–526. URL <https://www.aeaweb.org/articles?id=10.1257/app.20200367>.

- Branstetter, Lee, Neil Gandal, and Nadav Kuniesky. 2017. “Network-Mediated Knowledge Spillovers: A Cross-Country Comparative Analysis of Information Security Innovations.” Working Paper 23808, National Bureau of Economic Research. URL <http://www.nber.org/papers/w23808>.
- Claussen, Jörg, Oliver Falck, and Thorsten Grohsjean. 2012. “The strength of direct ties: Evidence from the electronic game industry.” *International Journal of Industrial Organization* 30 (2):223 – 230. URL <http://www.sciencedirect.com/science/article/pii/S016771871100083X>.
- Delcamp, Henry and Aija Leiponen. 2014. “Innovating standards through informal consortia: The case of wireless telecommunications.” *International Journal of Industrial Organization* 36:36–47. URL <https://www.sciencedirect.com/science/article/pii/S0167718713000787>. Standards, Intellectual Property & Innovation.

- Economides, Nicholas. 1996. "The economics of networks." *International Journal of Industrial Organization* 14 (6):673 – 699. URL <http://www.sciencedirect.com/science/article/pii/0167718796010156>.
- Economides, Nicholas and David Encaoua. 1996. "Special issue on network economics: Business conduct and market structure." *International Journal of Industrial Organization* 14 (6):669 – 671. URL <http://www.sciencedirect.com/science/article/pii/0167718796010193>.
- Fershtman, Chaim and Neil Gandal. 2011. "Direct and indirect knowledge spillovers: the “social network” of open-source projects." *The RAND Journal of Economics* 42 (1):70–91. URL <http://dx.doi.org/10.1111/j.1756-2171.2010.00126.x>.
- Gandal, Neil and Uriel Stettner. 2016. "Network dynamics and knowledge transfer in virtual organisations." *International Journal of Industrial Organization* 48:270–290. URL <https://www.sciencedirect.com/science/article/pii/S0167718716301278>.

- Ganguli, Ina, Jeffrey Lin, and Nicholas Reynolds. 2020. "The Paper Trail of Knowledge Spillovers: Evidence from Patent Interferences." *American Economic Journal: Applied Economics* 12 (2):278–302. URL <https://www.aeaweb.org/articles?id=10.1257/app.20180017>.
- Goyal, Sanjeev. 2012. *Connections: an introduction to the economics of networks*. Princeton University Press. URL <http://site.ebrary.com/lib/ubc/docDetail.action?docID=10364776>.
- Jackson, Matthew O. 2010. *Social and economic networks*. Princeton University Press.
- Malikov, Emir and Shunan Zhao. 2023. "On the Estimation of Cross-Firm Productivity Spillovers with an Application to FDI." *The Review of Economics and Statistics* 105 (5):1207–1223. URL https://doi.org/10.1162/rest_a_01080.

- Myers, Kyle R. and Lauren Lanahan. 2022. "Estimating Spillovers from Publicly Funded R&D: Evidence from the US Department of Energy." *American Economic Review* 112 (7):2393–2423. URL <https://www.aeaweb.org/articles?id=10.1257/aer.20210678>.
- Schnitzer, Monika and Martin Watzinger. 2022. "Measuring the Spillovers of Venture Capital." *The Review of Economics and Statistics* 104 (2):276–292. URL https://doi.org/10.1162/rest_a_00937.
- Zacchia, Paolo. 2019. "Knowledge Spillovers through Networks of Scientists." *The Review of Economic Studies* 87 (4):1989–2018. URL <https://doi.org/10.1093/restud/rdz033>.