Some *interesting* examples of field experiments with social networks.

Sharique Hasan
Fuqua School of Business, Duke University

Field: *Real* people; real meaningful *outcomes*.

Experiment: Randomization into conditions, done by experimenter or naturally in field setting.

There are hundreds of questions that we can try to answer about networks using field experiments.

I will focus on **three.**

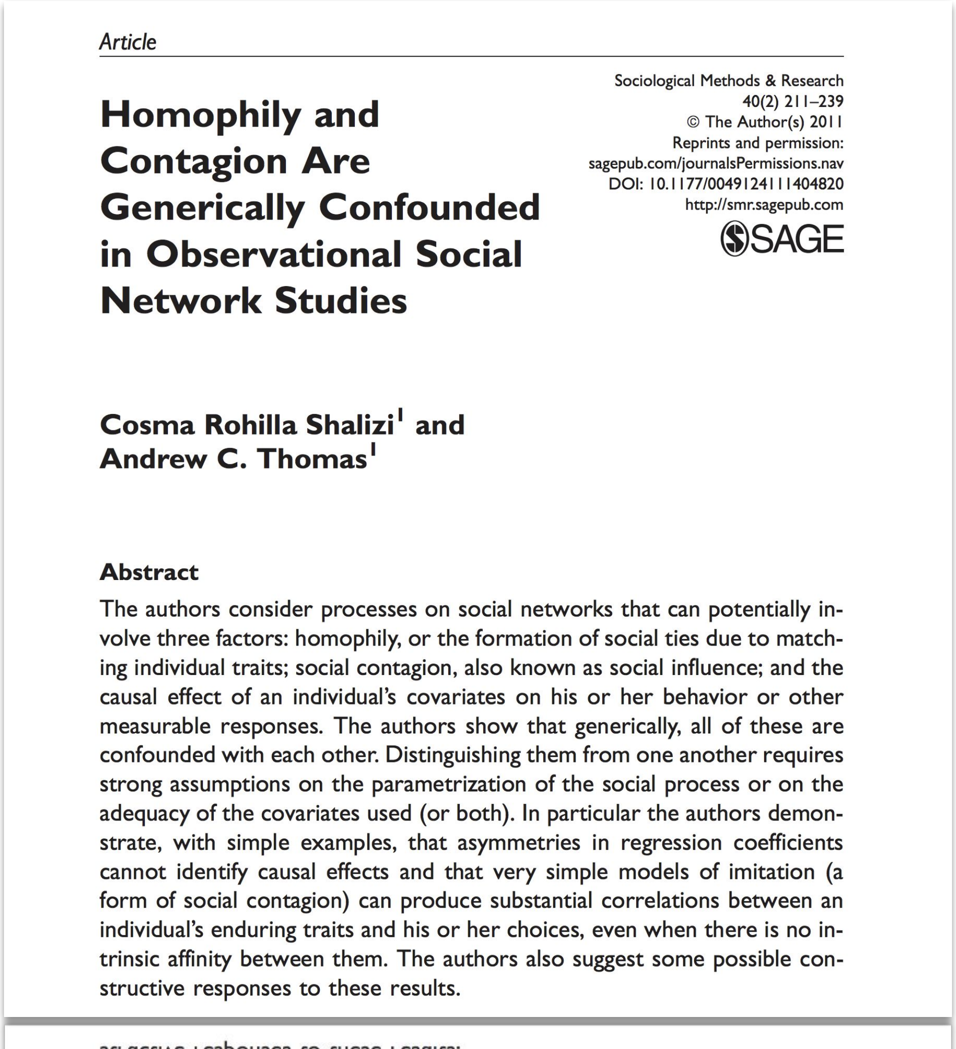
Peer effects: Does **j** influence the behavior/outcomes of ***i***?

Network Formation: What affects whether **i** forms a network tie with ***j***?

Designing networks: Which network structures maximize network-level outcomes?

There are hundreds of questions that we can try to answer about networks using field experiments.

Peer effects: Does **j** influence the behavior/outcomes of ***i***?



Article

Sociological Methods & Research
40(2) 211–239

DOI: 10.1177/0049124111404820
http://smr.sagepub.com
SAGE

Homophily and Contagion Are Generically Confounded in Observational Social Network Studies

Cosma Rohilla Shalizi[1] **and Andrew C. Thomas**[1]

Abstract

The authors consider processes on social networks that can potentially involve three factors: homophily, or the formation of social ties due to matching individual traits; social contagion, also known as social influence; and the causal effect of an individual's covariates on his or her behavior or other measurable responses. The authors show that generically, all of these are confounded with each other. Distinguishing them from one another requires strong assumptions on the parametrization of the social process or on the adequacy of the covariates used (or both). In particular the authors demonstrate, with simple examples, that asymmetries in regression coefficients cannot identify causal effects and that very simple models of imitation (a form of social contagion) can produce substantial correlations between an individual's enduring traits and his or her choices, even when there is no intrinsic affinity between them. The authors also suggest some possible constructive responses to these results.

Alice has good grades.

By interacting with Alice, will Bob also get good grades?

A	B	A's score	B's score
Alice	Bob	90	80
Bob	Alice	80	90
Raj	Sam	70	75
Sam	Raj	75	70

$$Y(i) = b0 + \mathbf{b1} * (Y(j)) + e$$

There are at least three problems with this approach:

Reflection: The regression specifies the influence of A on B.
We are picking up the influence of B on A.

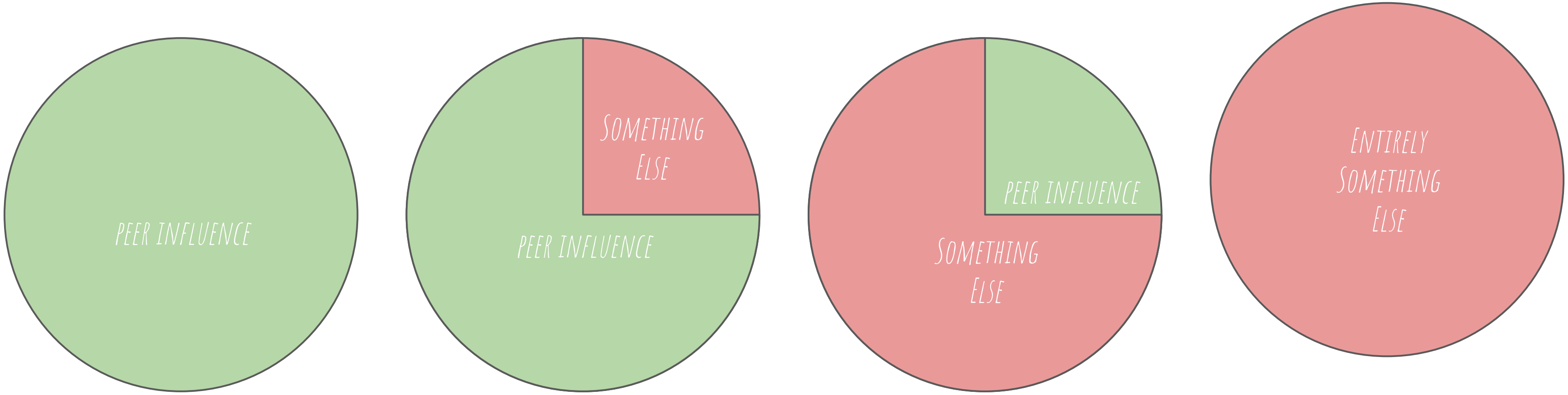
Selection: Similar people are more likely to interact.
We are picking up homophily, not influence.

Common shocks: People who are interacting have common contexts and experiences.
We are picking up the effect of common experiences (e.g., noisy hallways or extra coffee on the dorm floor).

*In essence, there is **imbalance between the treatment and control groups.***

The identification problem:

How much of **b1** is *actually* peer influence?



The standard approach to dealing with this identification problem:

Reflection: The regression specifies the influence of A on B.

Lag the independent variable.

Selection: Similar people are more likely to interact.

Include controls that account for dimensions on which people may decide to form a connection.

Common shocks: People who are interacting have common contexts and experiences.

Include features of the context that may influence the performance of both A and B.

$$Y(i,t) = b0 + \underline{\mathbf{b1}}*(Y(j,t-1)) + Controls + e$$

This seems like a reasonable strategy, ***but***:

How many other factors can you actually account for with "controls?"

Why randomization?:

1. Creates balance between treatment and control groups.
2. Reduces the likelihood that unobserved factors of selection or context are causing the observed effects.
3. This is because, treatment condition and the characteristics of pairs are now uncorrelated.

$$Y(i,t) = b0 + \underline{\mathbf{b1}}*(Y(j,t-1)) + e$$

PEER EFFECTS WITH RANDOM ASSIGNMENT: RESULTS FOR DARTMOUTH ROOMMATES*

BRUCE SACERDOTE

This paper uses a unique data set to measure peer effects among college roommates. Freshman year roommates and dormmates are randomly assigned at Dartmouth College. I find that peers have an impact on grade point average and on decisions to join social groups such as fraternities. Residential peer effects are markedly absent in other major life decisions such as choice of college major. Peer effects in GPA occur at the individual room level, whereas peer effects in fraternity membership occur both at the room level and the entire dorm level. Overall, the data provide strong evidence for the existence of peer effects in student outcomes.

I. INTRODUCTION

People have long believed that peer quality and behavior are among the most important determinants of student outcomes. This idea is expressed in the Coleman Report [1966], in Supreme Court decisions such as Brown versus Topeka Board of Education (1954), and in the findings of numerous researchers. Betts and Morell [1999] find that high school peer group characteristics affect undergraduate grade point average (GPA). Case and Katz [1991] find large peer effects on youth criminal behavior and drug use.[1] In a summary of the developmental psychology literature, Harris [1998] claims that parental behavior has no direct effect on child outcomes and that peer effects are the only important environmental factors affecting outcomes. A rich literature on neighborhood effects including Jencks and Mayer [1990], Rosenbaum [1992], and Katz, Kling, and Liebman [2001] shows that neighborhood peers can have profound effects on both adults and children.

Alice has good grades.

By interacting with Alice, will Bob also get good grades?

$$Y(i,t) = b0 + \underline{\mathbf{b1}}*(Y(j,t-1)) + Controls + e$$

TABLE III
PEER EFFECTS IN ACADEMIC OUTCOMES

	(1) Fresh year GPA	(2) Fresh year GPA w/ dorm f.e.	(3) Senior year GPA	(4) Fresh year GPA	(5) Fresh year GPA	(6) Fresh year GPA	(7) Graduate late	(8) Econ major
Roommates' GPA	0.120**	0.068**	0.008					
	(0.039)	(0.029)	(0.026)					
HS academic score (self)	0.014**	0.015**	0.013**				−0.0001	0.003**
	(0.0008)	(0.0007)	(0.0009)				(0.0003)	(0.0006)
HS academic score (roommates')	−0.001	−0.0003	0.0009				0.0003	−0.0001
	(0.001)	(0.0009)	(0.001)				(0.0003)	(0.0006)
roommates' academic score bottom 25 percent				0.016	0.014	0.017		
				(0.028)	(0.025)	(0.025)		
roommates' academic score top 25 percent				0.060**	0.047*	0.043*		
				(0.028)	(0.026)	(0.026)		
roommates' intention to graduate w/honors (1–4)						0.082**		
						(0.037)		
own academic score bottom 25 percent					−0.284**	−0.282**		
					(0.025)	(0.025)		

No evidence for a causal peer effect.
The coefficient is basically "0."

Check for updates

AMERICAN SOCIOLOGICAL ASSOCIATION

American Sociological Review
78(6) 1009–1032

DOI: 10.1177/0003122413505198
http://asr.sagepub.com
SAGE

The Mechanics of Social Capital and Academic Performance in an Indian College

Sharique Hasan[a] and Surendrakumar Bagde[b]

Abstract

In this article we examine how social capital affects the creation of human capital. Specifically, we study how college students' peers affect academic performance. Building on existing research, we consider the different types of peers in the academic context and the various mechanisms through which peers affect performance. We test our model using data from an engineering college in India. Our data include information about the performance of individual students as well as their randomly assigned roommates, chosen friends, and chosen study-partners. We find that students with able roommates perform better, and the magnitude of this roommate effect increases when the roommate's skills match the student's academic goals. We also find that students benefit equally from same- and different-caste roommates, suggesting that social similarity does not strengthen peer effects. Finally, although we do not find strong evidence for independent friendship or study-partner effects, our results suggest that roommates become study-partners, and in so doing, affect performance. Taken together, our findings demonstrate that peer effects are a consequential determinant of academic achievement.

Keywords

social capital, peer effects, social networks, learning, India

The metaphor of *social capital* has been an indispensable lens for scholars studying a diverse range of social and economic phe-

Indeed, social capital may have its greatest impact in the accumulation of *human capital*—the skills and knowledge that allow individu-

TABLE V
PEER EFFECTS IN SOCIAL OUTCOMES

	(1) Member frat/ soror	(2) Member frat/ soror	(3) Member frat/ soror	(4) Varsity athlete
roommate member of fraternity/sorority/coed	0.078**	0.056		
	(0.038)	(0.037)		
dorm average of fraternity/sorority/coed		0.321**		
		(0.135)		
roommate varsity athlete				0.045
				(0.033)
HS academic score (self)	0.0098	0.0011	0.0010	−0.004**
	(0.0010)	(0.0011)	(0.0011)	(0.001)
HS academic score (roommates')	−0.0017	−0.0016	−0.0016	−0.0002
	(0.0011)	(0.0011)	(0.0011)	(0.0007)
Own use of beer in high school (0–1)			0.135**	
			(0.038)	
Roommates' use of beer in high school (0–1)			−0.025	
			(0.026)	
			(0.026)	
Dormmates' use of beer in high school (0–1)			0.287**	
			(0.146)	
Dummies for housing questions	yes	yes	yes	yes
R^2	.02	.02	.03	.05
N	1589	1589	1589	1589

Standard errors are in parentheses and are corrected for clustering at the room level. In cases with more than one roommate, roommate variables are averaged. ** = p-value < .05.

Columns (1)–(4) are Probits. $\partial y/\partial x$ is shown.

In regression (2), dorm average of frat membership excludes own observation, and standard errors are corrected for clustering at dorm level.

In regression (3), use of beer in past year is coded 0–1 as follows: 0 = not at all, occasionally or frequently = 1. Dorm use of beer excludes own room and standard errors are corrected for clustering at dorm level.

Much of this was driven by two thing: (1) subject specific learning; (2) **peer effects were driven primarily by students who actually studied together.**

LETTER

doi:10.1038/nature11421

A 61-million-person experiment in social influence and political mobilization

Robert M. Bond[1], Christopher J. Fariss[1], Jason J. Jones[2], Adam D. I. Kramer[3], Cameron Marlow[3], Jaime E. Settle[1] & James H. Fowler[1,4]

Human behaviour is thought to spread through face-to-face social networks, but it is difficult to identify social influence effects in observational studies[9–13], and it is unknown whether online social networks operate in the same way[14–19]. Here we report results from a randomized controlled trial of political mobilization messages delivered to 61 million Facebook users during the 2010 US congressional elections. The results show that the messages directly influenced political self-expression, information seeking and real-world voting behaviour of millions of people. Furthermore, the messages not only influenced the users who received them but also the users' friends, and friends of friends. The effect of social transmission on real-world voting was greater than the direct effect of the messages themselves, and nearly all the transmission occurred between 'close friends' who were more likely to have a face-to-face relationship. These results suggest that strong ties are instrumental for spreading both online and real-world behaviour in human social networks.

Recent experimental studies[6,14–16] have attempted to measure the causal effect of social influence online. At the same time, there is increasing interest in the ability to use online social networks to study and influence real-world behaviour[17–19]. However, online social networks are also made up of many 'weak-tie' relationships[20] that may not facilitate social influence[21], and some studies suggest that online communication may not be an effective medium for influence[22]. An open question is whether online networks, which harness social information from face-to-face networks, can be used effectively to increase the likelihood of behaviour change and social contagion.

One behaviour that has been proposed to spread through networks is the act of voting in national elections. Voter turnout is significantly correlated among friends, family members and co-workers in observational studies[23,24]. Voter mobilization efforts are effective at increasing turnout[25], particularly those conducted face-to-face and those that appeal to social pressure[26] and social identity[27]. There is also evidence from one face-to-face field experiment that voting is 'contagious', in the sense that mobilization can spread from person to person within two-person households[28]. Although anecdotal accounts suggest that online mobilization has made a big difference in recent elections[21], a meta-analysis of email experiments suggests that online appeals to vote are ineffective[24].

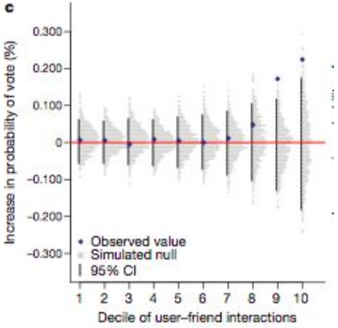
Voter mobilization experiments[26–28] have shown that most methods of contacting potential voters have small effects (if any) on turnout rates, ranging from 1% to 10%. However, the ability to reach large populations online means that even small effects could yield behaviour changes for millions of people. Furthermore, as many elections are competitive, these changes could affect electoral outcomes. For with all users of at least 18 years of age in the United States who accessed the Facebook website on 2 November 2010, the day of the US congressional elections. Users were randomly assigned to a 'social message' group, an 'informational message' group or a control group. The social message group ($n = 60{,}055{,}176$) was shown a statement at the top of their 'News Feed'. This message encouraged the user to vote, provided a link to find local polling places, showed a clickable button reading 'I Voted', showed a counter indicating how many other Facebook users had previously reported voting, and displayed up to six small randomly selected 'profile pictures' of the user's Facebook friends who had already clicked the I Voted button (Fig. 1). The informational message group ($n = 611{,}044$) was shown the message, poll information, counter and button, but they were not shown any faces of friends. The control group ($n = 613{,}096$) did not receive any message at the top of their News Feed.

The design of the experiment allowed us to assess the impact that the treatments had on three user actions; clicking the I Voted button, clicking the polling-place link and voting in the election. Clicking the I Voted button is similar to traditional measures of self-reported voting, but here users reported their vote to their social community rather than to a researcher. We therefore use this action to measure political self-expression, as it is likely to be affected by the extent to which a user desires to be seen as a voter by others. In contrast, social desirability should not affect other user actions in the same way. Clicking the polling-place link took users to a separate website that helped them to find a polling location, and this action was not reported to the user's social community. We therefore use this action to measure a user's desire to seek information about the election. Finally, we used a group-level process to study the validated voting behaviour of 6.3 million users matched to publicly available voter records (see Supplementary Information).

We first analyse direct effects. We cannot compare the treatment groups with the control group to assess the effect of the treatment on self-expression and information seeking, because the control group did not have the option to click an I Voted button or click on a polling-place link. However, we can compare the proportion of users between the two treatment groups to estimate the causal effect of seeing the faces of friends who have identified themselves as voters (Fig. 1). Users who received the social message were 2.08% (s.e.m., 0.05%; t-test, $P < 0.01$) more likely to click on the I Voted button than those who received the informational message (20.04% in the social message group versus 17.96% in the informational message group). Users who received the social message were also 0.26% (s.e.m., 0.02%; $P < 0.01$) more likely to click the polling-place information link than users who received the informational message (Fig. 1).

I encourage Alice to vote by showing an online message.

Does Bob, who is connected to Alice online, also vote?



"the messages not only influenced the users who received them but also the users' friends, and friends of friends. The effect of social transmission on real-world voting was greater than the direct effect of the messages themselves, and **nearly all the transmission occurred between 'close friends' who were more likely to have a face-to-face relationship. These results suggest that strong ties are instrumental for spreading both online and real-world behaviour in human social networks**"

MANAGEMENT SCIENCE
Vol. 61, No. 8, August 2015, pp. 1902–1920
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

informs
http://dx.doi.org/10.1287/mnsc.2014.2081

Do Your Online Friends Make You Pay? A Randomized Field Experiment on Peer Influence in Online Social Networks

Ravi Bapna, Akhmed Umyarov
Carlson School of Management, University of Minnesota, Minneapolis, Minnesota 55455
{rbapna@umn.edu, aumyarov@umn.edu}

Demonstrating compelling causal evidence of the existence and strength of peer-to-peer influence has become the holy grail of modern research in online social networks. In these networks, it has been consistently demonstrated that user characteristics and behavior tend to cluster both in space and in time. There are multiple well-known rival mechanisms that compete to be the explanation for this observed clustering. These range from peer influence to homophily to other unobservable external stimuli. These multiple mechanisms lead to similar observational data, yet have vastly different policy implications. In this paper, we present a novel randomized experiment that tests the existence of causal peer influence in the general population—one that did not involve subject recruitment for experimentation—of a particular large-scale online social network. We utilize a unique social feature to exogenously induce adoption of a paid service among a group of randomly selected users, and in the process develop a clean exogenous randomization of treatment and control groups. A variety of nonparametric, semiparametric, and parametric approaches, ranging from resampling-based inference to ego-level random effects to logistic regression to survival models, yield close to identical, statistically and economically significant estimates of peer influence in the general population of a freemium social network. Our estimates show that peer influence causes more than a 60% increase in odds of buying the service due to the influence coming from an adopting friend. In addition, we find that users with a smaller number of friends experience stronger relative increase in the adoption likelihood due to influence from their peers as compared to the users with a larger number of friends. Our nonparametric resampling procedure-based estimates are helpful in situations of networked data that violate independence assumptions. We establish that peer influence is a powerful force in getting users from free to premium levels, a known challenge in freemium communities.

Keywords: peer effects; randomized experiment; social contagion; nonparametric inference; freemium communities; online social networks
History: Received November 12, 2012; accepted July 12, 2014, by Sandra Slaughter, information systems. Published online in *Articles in Advance* April 8, 2015.

1. Introduction and Background

The general challenge of demonstrating causal inference from observational data has been immortalized in Manski's (1995) reference to the simultaneous movements of a man and his image in the mirror. He asks, "does the mirror cause the man's movement or reflect them?" (p. 1) and concludes that without understanding optics and human behavior, we cannot really tell. Interestingly, this quote from the pre-Facebook era is extremely relevant to the causality questions that arise in today's digital age. The growth of online social networks and the wide availability of online data has renewed interest in the identification of whether influence is "at play" in the general population of users of such networks. Today, a billion plus global citizens are socially connected by general networks such as Facebook and Twitter, as well as by niche networks such as Last.fm, Spotify, and LinkedIn among others.[1] These online social networks are credited with playing roles that range from inspiring political action to driving viral and word-of-mouth spread of products and services (Aral and Walker 2011, Hill et al. 2006, Iyengar et al. 2011, Manchanda et al. 2008, Mayzlin 2006), and as such, represent a vast reservoir of social and economic influence. Central to tapping into this reservoir is the understanding of causal relationships that drive the spread of products, services, and information over these social networks—the central focus of this paper.

[1] Online social networks such as Facebook, with a billion users, and Twitter, with more than a billion users, are consuming an increasingly significant portion of our time and attention. A recent CMO study estimated that 2013 is the first year in which the amount of time spent on social media exceeded that spent on TV, and that Facebook gets one in eight minutes users spend on the desktop, and one in five on the mobile (see Miners and CMO Staff 2013).

I encourage Alice to buy an online product.

Does Bob, who is connected to Alice online, also buy it?

Our estimates show that peer influence causes more than a 60% increase in odds of buying the service due to the influence coming from an adopting friend. In addition, **we find that users with a smaller number of friends experience stronger relative increase in the adoption likelihood** due to influence from their peers as compared to the users with a larger number of friends.

When does Advice Impact Startup Performance?*

Aaron Chatterji
Duke Fuqua and NBER
ronnie@duke.edu

Solène M. Delecourt
Stanford GSB
solened@stanford.edu

Sharique Hasan
Duke Fuqua
sh424@duke.edu

Rembrand Koning
Harvard Business School
rem@hbs.edu

May 17, 2018

Abstract

We conducted a field experiment to evaluate the impact of peer advice on the growth and survival outcomes of 100 high-growth startups over two years. We randomly grouped founders into pairs and had them provide each other advice over two days on topics ranging from people management to growth strategies. Founders who received advice from peers who actively managed their employees—with regular meetings, goal setting, and feedback—grew 28% larger and were 10% less likely to fail as compared to those who got advice from peers with a passive people-management approach. However, entrepreneurs with MBAs or accelerator experience did not respond to the advice of either active or passive peers, suggesting that formal training can limit the spread of informal management advice from peers.

Figure 2: The impact of advice on startup growth.

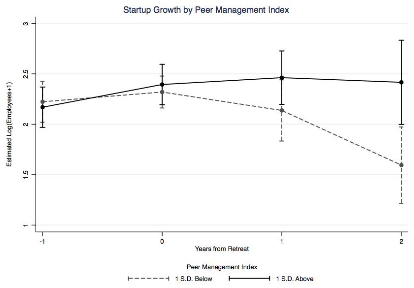
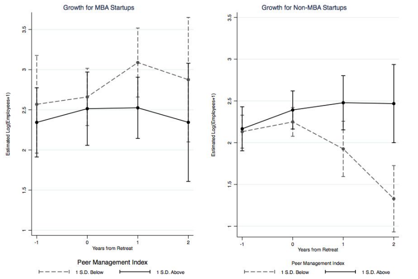


Figure 3: Estimated startup size by treatment status for startups with MBA-founders (23 Startups) and for non-MBA founders (67 Startups).



Founders who received advice from peers who actively managed their employees---with regular meetings, goal setting, and feedback---grew their firms to be 28% larger and were 10% less likely to fail as compared to those who got advice from peers with a passive people-management approach. However, entrepreneurs with MBAs or accelerator experience did not respond to the advice of either active or passive peers, suggesting that formal training can limit the spread of informal management advice from peers.

Network Formation: What affects whether **i** form a network tie with ***j***?

MANAGEMENT SCIENCE
Vol. 62, No. 11, November 2016, pp. 3100–3122
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

informs
https://doi.org/10.1287/mnsc.2015.2301

One-Way Mirrors in Online Dating: A Randomized Field Experiment

Ravi Bapna
Carlson School of Management, University of Minnesota, Minneapolis, Minnesota 55455, rbapna@umn.edu
Jui Ramaprasad
Desautels Faculty of Management, McGill University, Montréal, Quebec H3A 1G5, Canada, jui.ramaprasad@mcgill.ca
Galit Shmueli
Institute of Service Science, National Tsing Hua University, Hsinchu, Taiwan 30013, galit.shmueli@iss.nthu.edu.tw
Akhmed Umyarov
Carlson School of Management, University of Minnesota, Minneapolis, Minnesota 55455, aumyarov@umn.edu

The growing popularity of online dating websites is altering one of the most fundamental human activities: finding a date or a marriage partner. Online dating platforms offer new capabilities, such as extensive search, big data–based mate recommendations, and varying levels of anonymity, whose parallels do not exist in the physical world. Yet little is known about the causal effects of these new features. In this study we examine the impact of a particular anonymity feature, which is unique to online environments, on matching outcomes. This feature allows users to browse profiles of other users anonymously, by being able to check out a potential mate's profile while not leaving any visible online record of the visit. Although this feature may decrease search costs and allow users to search without inhibition, it also eliminates "weak signals" of interest for their potential mates that may play an important role in establishing successful communication. We run a randomized field experiment on a major North American online dating website, where 50,000 of 100,000 randomly selected new users are gifted the ability to anonymously view profiles of other users. Compared with the control group, the users treated with anonymity become disinhibited, in that they view more profiles and are more likely to view same-sex and interracial mates. However, based on our analysis, we demonstrate causally that weak signaling is a key mechanism in achieving higher levels of matching outcomes. Anonymous users, who lose the ability to leave a weak signal, end up having fewer matches compared with their nonanonymous counterparts. This effect of anonymity is particularly strong for women, who tend not to make the first move and instead rely on the counterparty to initiate the communication. Further, the reduction in quantity of matches by anonymous users is not compensated by a corresponding increase in quality of matches.

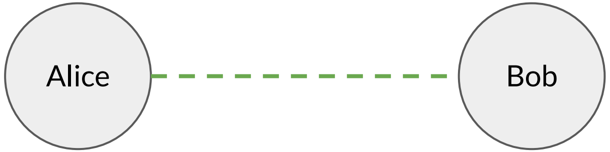
Keywords: online dating; anonymity; weak signaling; randomized trial; field experiment
History: Received July 29, 2013; accepted April 21, 2015, by Lorin Hitt, information systems. Published online in *Articles in Advance* February 2, 2016.

1. Motivation and Background

According to the U.S. Census Bureau, 46% of the single population in the United States uses online dating[1] to initiate and engage in the process of selecting a partner for reasons ranging from finding companionship to marrying and conceiving children and everything in between. Finding the optimal dating and ultimately marriage partner is one of the most important socioeconomic decisions made by humans. Yet such dating markets are fraught with frictions and inefficiencies, often leading people to rely on choices made through happenstance—an offhand referral, or perhaps a late night at the office (Paumgarten 2011). Interestingly, this primal human activity is being reshaped with the advent of big data and algorithmic matchmaking (Slater 2013). The continued growth of online dating despite the presence of a close substitute, the physical world, reflects the presence of significant frictions in the offline dating and marriage markets. Yet the underlying processes, dynamics, and implications of mate seeking in the online world are largely unstudied. Also unknown are the implications of the new features and capabilities that these new online matching markets bring to an age-old human activity. In this paper, we address this gap by studying the causal impact of anonymity, a key feature unique to the online environment, through a randomized experiment in partnership with a major online

[1] Of the 87 million singles in the United States, nearly half of them, or 40 million, have tried online dating, according to the U.S. Census Bureau (Gelles 2011).

What encourages Alice and Bob to form a tie?



"I was looking back to see if you were looking back at me to see me looking back at you"

We run a randomized field experiment on a major North American online dating website, where 50,000 of 100,000 randomly selected new users are gifted the ability to anonymously view profiles of other users. **Compared with the control group, the users treated with anonymity become disinhibited, in that they view more profiles and are more likely to view same-sex and interracial mates**.

Anonymous users, who lose the ability to leave a weak signal, end up having fewer matches compared with their nonanonymous counterparts. This effect of anonymity is particularly strong for women, who tend not to make the first move and instead rely on the counterparty to initiate the communication.

The Review *of* Economics *and* Statistics

Vol. XCIX — October 2017 — Number 4

A FIELD EXPERIMENT ON SEARCH COSTS AND THE FORMATION OF SCIENTIFIC COLLABORATIONS

Kevin J. Boudreau, Tom Brady, Ina Ganguli, Patrick Gaule, Eva Guinan, Anthony Hollenberg, and Karim R. Lakhani*

Abstract—We present the results of a field experiment conducted at Harvard Medical School to understand the extent to which search costs affect matching among scientific collaborators. We generated exogenous variation in search costs for pairs of potential collaborators by randomly assigning individuals to 90-minute structured information-sharing sessions as part of a grant funding opportunity. We estimate that the treatment increases the probability of grant co-application of a given pair of researchers by 75%. The findings suggest that matching between scientists is subject to considerable friction, even in the case of geographically proximate scientists working in the same institutional context.

I. Introduction

THE primary unit of scientific knowledge production has become the team or collaboration rather than the lone scientist (Jones, 2009). Indeed, teams are not only growing in frequency, but also in size and impact relative to single authors (Wuchty, Jones, & Uzzi, 2007). Unlike settings inside firms, where executives and managers play a central role in organizing and forming teams (Lazear & Shaw, 2007), academic scientists have greater freedom and autonomy in selecting their collaborators and their topics of inquiry (Stephan 2012). Although there is a growing body of research on the productivity and outcomes of scientific teams once formed (e.g., Adams et al., 2005; Wuchty, Jones, & Uzzi, 2007; Agrawal, Goldfarb, & Teodoridis, 2016), we know relatively little about the largely decentralized process by which scientific teams come into existence (Stephan, 2012). In this paper, we investigate the role of one particular mechanism, search costs and frictions, on these matching outcomes.

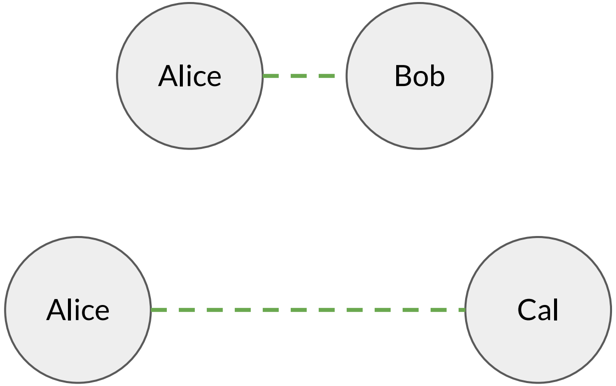
The role of search costs and resulting frictions in the formation of scientific collaborations is not well understood. On the one hand, the growing prominence of teams and falling communications and collaboration costs in science (Agrawal & Goldfarb, 2008; Ding et al., 2010) might suggest forces favorable to novel team formation. On the other hand, geography and distance are regularly documented to play a role in shaping collaborations, even today (e.g., Rosenthal & Strange, 2001; Glaeser, 2010, Catalini, 2016), and rather than continually forming novel collaborations, scientists most often work with partners in the same institution, in a similar knowledge domain, and within preexisting social networks (Baccara & Yariv, 2013; Freeman, Ganguli, & Murciano-Goroff, 2015; Freeman & Huang, 2014; Fafchamps, Goyal, & Van der Leij, 2010; Azoulay, Liu, & Stuart, 2009). Moreover, past collaborations remain an important predictor of future ones. Although these patterns might be explained by any number of factors, they raise the question of whether search costs play a first-order role in shaping the organization of scientists into teams.

The high information requirements for forming matches suggest that search frictions may be an important consideration. A large number of factors can play a role in decisions to collaborate, and these factors may be nuanced or difficult to observe. This includes factors such as the complementar-

Received for publication March 11, 2016. Revision accepted for publication December 2, 2016. Editor: Brigitte C. Madrian.

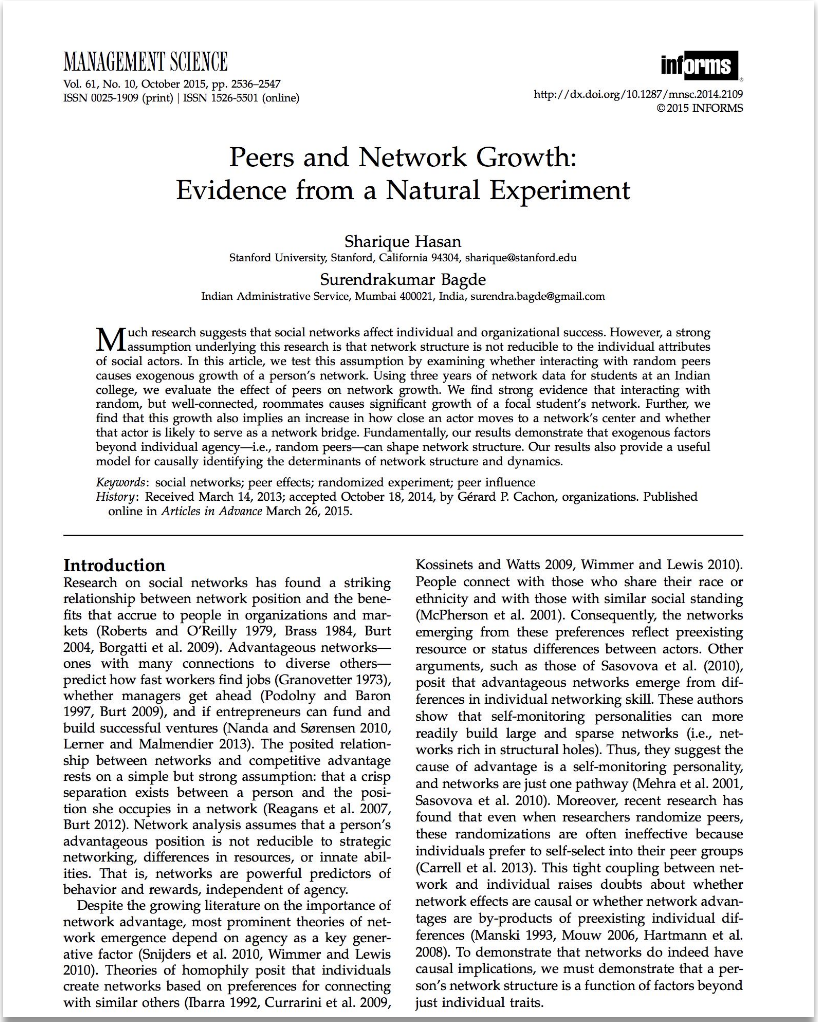
* Boudreau: Northeastern University; Brady: Massachusetts General Hospital and Harvard Medical School; Ganguli: University of Massachusetts Amherst; Gaule: CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Academy of Sciences of the Czech Republic; Guinan: Dana-Farber Cancer Institute and Harvard Medical School; Hollenberg: Beth Israel Deaconess Medical Center and Harvard Medical School; Lakhani: Harvard Business School and NBER.

K.J.B., I.G., P.G., E.G., and K.R.L. designed, developed, and executed the experiment and contributed to the manuscript. I.G. and P.G. conducted data analysis. T.B. and A.H. contributed to the experimental development and execution. We appreciate helpful comments from Pierre Azoulay, Marcel Fafchamps, Lee Fleming, Ben Golub, Shane Greenstein, Ben Jones, Nicola Lacetera, Josh Lerner, Paula Stephan, and Scott Stern, and seminar participants at the NBER Summer Institute Innovation Meetings, the NBER Productivity Lunch, the Georgia Tech REER Conference, MIT Sloan School of Management, Northwestern University, University of California, Berkeley, University of Rochester, SITE Stockholm School of Economics, CIRCLE Lund University, Graduate Institute Geneva, and Universidad Carlos III de Madrid. This work was conducted with support from Harvard Catalyst/The Harvard Clinical and Translational Science Center (National Center for Advancing Translational Sciences, National Institutes of Health Awards UL1TR001102, UL1TR000170, and UL1RR025758-02S4), NASA Tournament Lab at Harvard University, Harvard Business School Division of Research and Faculty Development, and financial contributions from Harvard University and its affiliated academic health care centers. We also thank Harvard Catalyst for support and cooperation in implementing the experiment, particularly Lee Nadler, William Chin, Laura Weisel, and David Frank. Amy Webber and Wei Zhong provided valuable assistance throughout the implementation pro-



Vary search costs for pairs of potential collaborators by randomly assigning individuals to 90-minute structured information-sharing sessions as part of a grant funding opportunity.

We estimate that the treatment **increases the probability of grant co-application of a given pair of researchers by 75%.** The findings suggest that matching between scientists is subject to considerable friction, ***even in the case of geographically proximate scientists working in the same institutional context.***


MANAGEMENT SCIENCE
Vol. 61, No. 10, October 2015, pp. 2536–2547
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

informs
http://dx.doi.org/10.1287/mnsc.2014.2109

Peers and Network Growth: Evidence from a Natural Experiment

Sharique Hasan
Stanford University, Stanford, California 94304, sharique@stanford.edu

Surendrakumar Bagde
Indian Administrative Service, Mumbai 400021, India, surendra.bagde@gmail.com

Much research suggests that social networks affect individual and organizational success. However, a strong assumption underlying this research is that network structure is not reducible to the individual attributes of social actors. In this article, we test this assumption by examining whether interacting with random peers causes exogenous growth of a person's network. Using three years of network data for students at an Indian college, we evaluate the effect of peers on network growth. We find strong evidence that interacting with random, but well-connected, roommates causes significant growth of a focal student's network. Further, we find that this growth also implies an increase in how close an actor moves to a network's center and whether that actor is likely to serve as a network bridge. Fundamentally, our results demonstrate that exogenous factors beyond individual agency—i.e., random peers—can shape network structure. Our results also provide a useful model for causally identifying the determinants of network structure and dynamics.

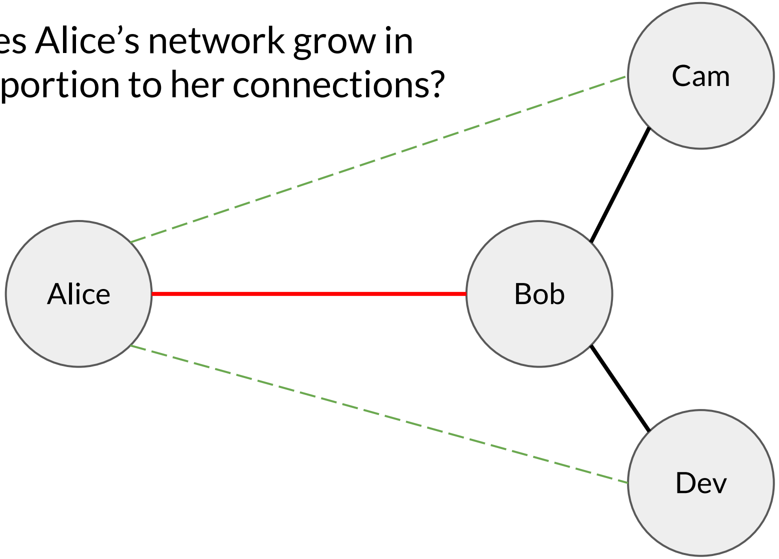
Keywords: social networks; peer effects; randomized experiment; peer influence
History: Received March 14, 2013; accepted October 18, 2014, by Gérard P. Cachon, organizations. Published online in *Articles in Advance* March 26, 2015.

Introduction

Research on social networks has found a striking relationship between network position and the benefits that accrue to people in organizations and markets (Roberts and O'Reilly 1979, Brass 1984, Burt 2004, Borgatti et al. 2009). Advantageous networks—ones with many connections to diverse others—predict how fast workers find jobs (Granovetter 1973), whether managers get ahead (Podolny and Baron 1997, Burt 2009), and if entrepreneurs can fund and build successful ventures (Nanda and Sørensen 2010, Lerner and Malmendier 2013). The posited relationship between networks and competitive advantage rests on a simple but strong assumption: that a crisp separation exists between a person and the position she occupies in a network (Reagans et al. 2007, Burt 2012). Network analysis assumes that a person's advantageous position is not reducible to strategic networking, differences in resources, or innate abilities. That is, networks are powerful predictors of behavior and rewards, independent of agency.

Despite the growing literature on the importance of network advantage, most prominent theories of network emergence depend on agency as a key generative factor (Snijders et al. 2010, Wimmer and Lewis 2010). Theories of homophily posit that individuals create networks based on preferences for connecting with similar others (Ibarra 1992, Currarini et al. 2009, Kossinets and Watts 2009, Wimmer and Lewis 2010). People connect with those who share their race or ethnicity and with those with similar social standing (McPherson et al. 2001). Consequently, the networks emerging from these preferences reflect preexisting resource or status differences between actors. Other arguments, such as those of Sasovova et al. (2010), posit that advantageous networks emerge from differences in individual networking skill. These authors show that self-monitoring personalities can more readily build large and sparse networks (i.e., networks rich in structural holes). Thus, they suggest the cause of advantage is a self-monitoring personality, and networks are just one pathway (Mehra et al. 2001, Sasovova et al. 2010). Moreover, recent research has found that even when researchers randomize peers, these randomizations are often ineffective because individuals prefer to self-select into their peer groups (Carrell et al. 2013). This tight coupling between network and individual raises doubts about whether network effects are causal or whether network advantages are by-products of preexisting individual differences (Manski 1993, Mouw 2006, Hartmann et al. 2008). To demonstrate that networks do indeed have causal implications, we must demonstrate that a person's network structure is a function of factors beyond just individual traits.

Does Alice's network grow in proportion to her connections?



We find strong evidence that interacting with random, but well-connected, roommates causes significant growth of a focal student's network. Further, we find that this growth also implies an increase in how close an actor moves to a network's center and whether that actor is likely to serve as a network bridge.

Designing networks: Which network structures maximize network-level outcomes?

The Spread of Behavior in an Online Social Network Experiment

Damon Centola

How do social networks affect the spread of behavior? A popular hypothesis states that networks with many clustered ties and a high degree of separation will be less effective for behavioral diffusion than networks in which locally redundant ties are rewired to provide shortcuts across the social space. A competing hypothesis argues that when behaviors require social reinforcement, a network with more clustering may be more advantageous, even if the network as a whole has a larger diameter. I investigated the effects of network structure on diffusion by studying the spread of health behavior through artificially structured online communities. Individual adoption was much more likely when participants received social reinforcement from multiple neighbors in the social network. The behavior spread farther and faster across clustered-lattice networks than across corresponding random networks.

Many behaviors spread through social contact (1–3). As a result, the network structure of who is connected to whom can critically affect the extent to which a behavior diffuses across a population (2–8). There are two competing hypotheses about how network structure affects diffusion. The "strength of weak ties" hypothesis predicts that networks with many "long ties" (e.g., "small-world" topologies) will spread a social behavior farther and more quickly than a network in which ties are highly clustered (4–6). This hypothesis treats the spread of behavior as a simple contagion, such as disease or information: A single contact with an "infected" individual is usually sufficient to transmit the behavior (2). The power of long ties is that they reduce the redundancy of the diffusion process by connecting people whose friends do not know each other, thereby allowing a behavior to rapidly spread to other areas of the network (3–5). The ideal case for this lack of redundancy is a "random" network, in which, in expectation for a large population, each of an individual's ties reaches out to different neighborhoods (4, 9). The other hypothesis states that, unlike disease, social behavior is a complex contagion: People usually require contact with multiple sources of "infection" before being convinced to adopt a behavior (2). This hypothesis predicts that because clustered networks have more redundant ties, which provide social reinforcement for adoption, they will better promote the diffusion of behaviors across large populations (2, 7). Despite the scientific (6, 7, 10) and practical (1, 2, 11) importance of understanding the spread of behavior through social networks, an empirical test of these predictions has not been possible, because it requires the ability to independently vary the topological structure of a social network (12).

I tested the effects of network structure on diffusion using a controlled experimental approach. I studied the spread of a health behavior through a network-embedded population by creating an Internet-based health community, containing 1528 participants recruited from health-interest World Wide Web sites (13).

Each participant created an anonymous online profile, including an avatar, a user name, and a set of health interests. They were then matched with other participants in the study—referred to as "health buddies"—as members of an online health community. Participants could not contact their health buddies directly, but they could receive emails from the study informing them of their health buddies' activities. To preserve anonymity and to prevent people from trying to identify friends who may have also signed up for the study (or from trying to contact health buddies outside the context of the experiment), I blinded the identifiers that people used. Participants made decisions about whether or not to adopt a health behavior based on the adoption patterns of their health buddies. The health behavior used for this study was the decision to register for an Internet-based health forum, which offered access and rating tools for online health resources (13).

The health forum was not known (or accessible) to anyone except participants in the experiment. This ensured that the only sources of encouragement that participants had to join the forum were the signals that they received from their health buddies. The forum was populated with initial ratings to provide content for the early adopters. However, all subsequent content was contributed by the participants who joined the forum.

Participants arriving to the study were randomly assigned to one of two experimental conditions—a clustered-lattice network and a random network—that were distinguished only by the topological structure of the social networks (Fig. 1). In the clustered-lattice–network condition, there was a high level of clustering (5, 6, 13) created by redundant ties that linked each node's neighbors to one another. The random network condition was created by rewiring the clustered-lattice network via a permutation algorithm based on the small-world–network model (6, 13–15). This ensured that each node maintained the exact same number of neighbors as in the clustered network (that is, a homogeneous degree distribution), while simultaneously reducing clustering in the network and eliminating redundant ties within and between neighborhoods (4, 6, 14).

The network topologies were created before the participants arrived, and the participants could

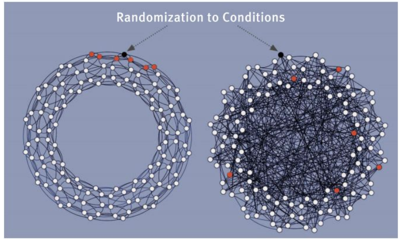
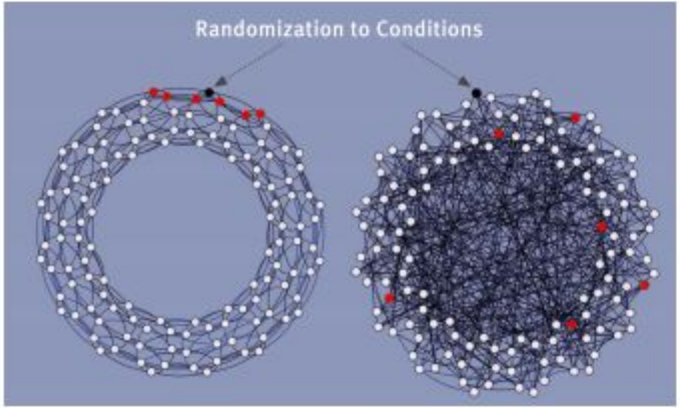


Fig. 1. Randomization of participants to clustered-lattice and random-network conditions in a single trial of this study (N = 128, Z = 6). In each condition, the black node shows the focal node of a neighborhood to which an individual is being assigned, and the red nodes correspond to that individual's neighbors in the network. In the clustered-lattice network, the red nodes share neighbors with each other, whereas in the random network they do not. White nodes indicate individuals who are not connected to the focal node.

Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02142, USA. E-mail: dcentola@mit.edu

Fig. 1. Randomization of participants to clustered-lattice and random-network conditions in a single trial of this study (N = 128, Z = 6). In each condition, the black node shows the focal node of a neighborhood to which an individual is being assigned, and the red nodes correspond to that individual's neighbors in the network. In the clustered-lattice network, the red nodes share neighbors with each other, whereas in the random network they do not. White nodes indicate individuals who are not connected to the focal node.



"behavior spread farther and faster across clustered-lattice networks than across corresponding random networks."

Can Network Theory-based Targeting Increase Technology Adoption?

Lori Beaman — Northwestern Univ.
Ariel BenYishay — Coll. of William and Mary
Jeremy Magruder — UC-Berkeley
Ahmed Mushfiq Mobarak — Yale University

June 2015

Abstract

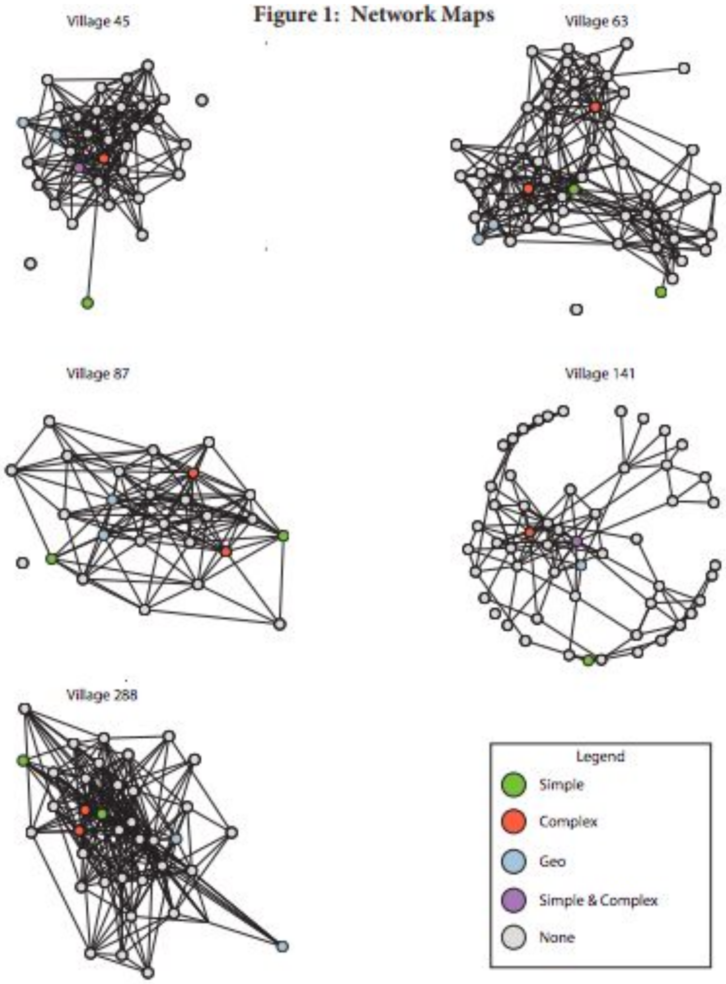
In order to induce farmers to adopt a productive new agricultural technology, we apply simple and complex contagion diffusion models on rich social network data from 200 villages in Malawi to identify seed farmers to target and train on the new technology. A randomized controlled trial compares these theory-driven network targeting approaches to simpler strategies that either rely on a government extension worker or an easily measurable proxy for the social network (geographic distance between households) to identify seed farmers. Both reduced form and structural estimates suggest a learning environment in which most farmers need to learn about the technology from multiple people before they adopt themselves.

JEL Codes: O16, O13

Keywords: Social Learning, Agricultural Technology Adoption, Complex Contagion, Malawi

* Contact: Beaman: l-beaman@northwestern.edu, BenYishay: abenyishay@wm.edu, Magruder: jmagruder@berkeley.edu, Mobarak: ahmed.mobarak@yale.edu. We thank the CEGA/JPAL Agricultural Technology Adoption Initiative (ATAI) and 3ie for financial support. We gratefully acknowledge the support and cooperation of Paul Fatch, Readwell Musopole and many other staff members of the Malawi Ministry of Agriculture. Thomas Coen, Niall Kelleher, Maria Jones, Ofer Cohen, Allen Baumgardner-Zuzik, and the IPA-Malawi country office provided invaluable support for data collection. Hossein Alidaee and Tetyana Zelenska provided excellent research assistance. We thank, without implicating, Arun Chandrasekhar, Matt Jackson, Kaivan Munshi, Chris Udry and numerous seminar audiences for very helpful comments.

Figure 1: Network Maps



Can Network Theory-based Targeting Increase Technology Adoption?

Lori Beaman, Northwestern Univ.
Ariel BenYishay, Coll. of William and Mary
Jeremy Magruder, UC-Berkeley
Ahmed Mushfiq Mobarak, Yale University

June 2015

Abstract

In order to induce farmers to adopt a productive new agricultural technology, we apply simple and complex contagion diffusion models on rich social network data from 200 villages in Malawi to identify seed farmers to target and train on the new technology. A randomized controlled trial compares these theory-driven network targeting approaches to simpler strategies that either rely on a government extension worker or an easily measurable proxy for the social network (geographic distance between households) to identify seed farmers. Both reduced form and structural estimates suggest a learning environment in which most farmers need to learn about the technology from multiple people before they adopt themselves.

JEL Codes: O16, O13

Keywords: Social Learning, Agricultural Technology Adoption, Complex Contagion, Malawi

* Contact: Beaman: l-beaman@northwestern.edu, BenYishay: abenyishay@wm.edu, Magruder: jmagruder@berkeley.edu, Mobarak: ahmed.mobarak@yale.edu. We thank the CEGA/JPAL Agricultural Technology Adoption Initiative (ATAI) and 3ie for financial support. We gratefully acknowledge the support and cooperation of Paul Fatch, Readwell Musopole and many other staff members of the Malawi Ministry of Agriculture. Thomas Coen, Niall Kelleher, Maria Jones, Ofer Cohen, Allen Baumgardner-Zuzik, and the IPA-Malawi country office provided invaluable support for data collection. Hossein Alidaee and Tetyana Zelenska provided excellent research assistance. We thank, without implicating, Arun Chandrasekhar, Matt Jackson, Kaivan Munshi, Chris Udry and numerous seminar audiences for very helpful comments.

Table 6: Village-Level Regressions of Adoption Outcomes Across Treatment Arms

	Adoption Rate		Number Adopters		Any Non-Seed Adopters	
	(1)	(2)	(3)	(4)	(5)	(6)
Simple Contagion Treatment	0.036 **	0.006	1.119	0.283	0.155	0.189 *
	(0.017)	(0.022)	(0.859)	(1.561)	(0.100)	(0.111)
Complex Contagion Treatment	0.036 **	0.036	2.876 **	2.237	0.252 ***	0.304 ***
	(0.016)	(0.026)	(1.386)	(2.007)	(0.093)	(0.101)
Geographic treatment	0.038	0.013	0.356	-1.379	0.107	0.188 *
	(0.027)	(0.034)	(0.823)	(1.284)	(0.096)	(0.110)
Year	2	3	2	3	2	3
N	200	141	200	141	200	141
Mean of Benchmark Treatment (omitted category)	0.038	0.075	2.11	4.81	0.420	0.543
SD of Benchmark	0.073	0.109	4.17	7.81	0.499	0.505
p-values for equality in coefficients:						
Simple = Complex	0.981	0.173	0.258	0.373	0.300	0.240
Complex = Geo	0.937	0.491	0.084	0.061	0.102	0.220
Simple = Geo	0.950	0.783	0.424	0.237	0.623	0.990

In order to induce farmers to adopt a productive new agricultural technology, we apply simple and complex contagion diffusion models on rich social network data from 200 villages in Malawi to identify seed farmers to target and train on the new technology

Econometrica, Vol. 81, No. 3 (May, 2013), 855–882

FROM NATURAL VARIATION TO OPTIMAL POLICY? THE IMPORTANCE OF ENDOGENOUS PEER GROUP FORMATION

BY SCOTT E. CARRELL, BRUCE I. SACERDOTE, AND JAMES E. WEST[1]

We take cohorts of entering freshmen at the United States Air Force Academy and assign half to peer groups designed to maximize the academic performance of the lowest ability students. Our assignment algorithm uses nonlinear peer effects estimates from the historical pre-treatment data, in which students were randomly assigned to peer groups. We find a negative and significant treatment effect for the students we intended to help. We provide evidence that within our "optimally" designed peer groups, students avoided the peers with whom we intended them to interact and instead formed more homogeneous subgroups. These results illustrate how policies that manipulate peer groups for a desired social outcome can be confounded by changes in the endogenous patterns of social interactions within the group.

KEYWORDS: Peer effects, social network formation, homophily.

0. INTRODUCTION

PEER EFFECTS HAVE BEEN widely studied in the economics literature due to the perceived importance peers play in workplace, educational, and behavioral outcomes. Previous studies in the economics literature have focused almost exclusively on the *identification* of peer effects and have only hinted at the potential policy implications of the results.[2] Recent econometric studies on assortative matching by Graham, Imbens, and Ridder (2009) and Bhattacharya (2009) have theorized that individuals could be sorted into peer groups to maximize productivity.[3]

This study takes a first step in determining whether student academic performance can be improved through the systematic sorting of students into peer groups. We first identify nonlinear peer effects at the United States Air Force Academy (USAFA) using pre-treatment data in which students were randomly assigned to peer groups (squadrons) of about 30 students. These estimates showed that low ability students benefited significantly from being with peers who have high SAT Verbal scores. We use these estimates to create optimally designed peer groups intended to improve academic achievement of the

mally designed peer groups intended to improve academic achievement of the

Alice has good grades.

By interacting with Alice, will Bob also get good grades?

$$Y(i,t) = b0 + \underline{\mathbf{b1}}*(Y(j,t-1)) + \underline{\mathbf{b2}}*(Y(i,t-1)) + \underline{\mathbf{b3}}*(Y(j,t-1)*Y(i,t-1)) + e$$

We provide evidence that within our "optimally" designed peer groups, s**tudents avoided the peers with whom we intended them to interact and instead formed more homogeneous subgroups**.

These results illustrate how policies that manipulate peer groups for a desired social outcome can be confounded by changes in the endogenous patterns of social interactions within the group.

Three types of network field experiments**.**

Peer effects: Does **j** influence the behavior/outcomes of ***i***?

Network Formation: What affects whether **i** forms a network tie with ***j***?

Designing networks: Which network structures maximize network-level outcomes?