Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Peers and Network Growth: Evidence from a Natural Experiment

Sharique Hasan, Surendrakumar Bagde

To cite this article:
Sharique Hasan, Surendrakumar Bagde (2015) Peers and Network Growth: Evidence from a Natural Experiment. Management Science 61(10):2536-2547. http://dx.doi.org/10.1287/mnsc.2014.2109

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.

One pathway theorized to create exogenous divergence in network structure focuses on the interactions a person has with others—sometimes called peers—in shared organizational settings (Feld 1981, Small et al. 2008, Small 2010). Such settings include schools, where peers are other students; businesses, where peers are coworkers; and voluntary organizations, where peers are comembers. Through these interactions, an individual's network will come to mirror the opportunity structure represented by her peer group (Feld 1981, 1982). By incorporating her peers into her network, a focal individual's network can improve through peer interaction in less direct ways as well. Peers in organizational settings possess networks and connections of their own. These networks are a resource that helps a focal individual to grow and alter her own network.

A peer's network can be a catalyst for the focal individual in several ways. First, a peer can serve as an active connector who links the focal actor to her friends and acquaintances (Fernandez and Weinberg 1997, Obstfeld 2005). When a focal actor needs an introduction to someone possessing some skill or resource, a peer can make the introduction. A peer can also facilitate introductions by having the focal individual join her at professional and social gatherings where other members of the peer's network are present. By introducing the focal individual to members of her network, a peer facilitates closure—i.e., creates connections between two previously disconnected parties (Kossinets and Watts 2006). Through closure, a peer's connection becomes the focal individual's own, thereby growing the latter's network. However, some peers are more helpful than others in this process. Some peers have large networks; others have small ones. A well-connected peer—one with a large network and high "indegree"—can facilitate more introductions and thus aids the focal actor in growing her network more easily. Besides directly facilitating connections, a well-connected peer may catalyze network growth indirectly. First, a peer skilled in networking (as reflected by a large network) may teach those skills to the focal actor (Burt and Ronchi 2007). With these networking skills, the focal actor can independently grow her network, even without direct introductions. Second, a peer may introduce the focal actor to a setting or context (a club, party, etc.) where she meets individuals not already connected to the peer (Feld 1982). Third, an association with a popular peer may serve as a signal to others that the focal individual is a desirable friend or acquaintance (Podolny 2001, Gould 2002). All these mechanisms imply the same effect for the focal individual: having better-connected peers causes more network growth. Finally, adding new network ties may also cause an increase in how close a focal individual becomes to others in the larger network and whether she is more likely to act as a network bridge. The latter is particularly true if newly created ties are to nonredundant parts of the network.

Unfortunately, demonstrating that peers affect network growth—and thus that network position is not reducible to individual attributes—is difficult (Manski 1993, Hartmann et al. 2008, Aral and Walker 2011). Two challenges are particularly salient. The first challenge is called the *selection problem* and arises because peer groups are endogenous. Individuals select (or are selected into) organizational settings, and thus, the composition of their peer groups is still a by-product of individual traits and agency. For example, peer groups at elite MBA programs help students to grow and diversify their networks. However, admission into such program often requires above-average ability, networks, or status. The same issues arise when examining the network benefits of employment at a prestigious firm, membership in an exclusive club, or residence in an expensive neighborhood. Thus, selection confounds the peer effect on networks, and correlation in network size and structure between a peer and focal individual does not imply causation (Manski 1993). The second challenge is called the *reflection problem*. Reflection arises because the networks of the focal individual and peer are often jointly determined. Although we theorize that the networks of peers affect the focal individual's network, the opposite can equally hold. Further, an environmental characteristic (e.g., something special about the setting) can also simultaneously cause both the individual's and peer's networks to grow.

Therefore, identifying a peer's effect on network growth requires resolving the two challenges described above. To do so, we require data with two characteristics (Sacerdote 2001, Marmaros and Sacerdote 2006). To address selection, focal individuals must be matched to their peers so that peer and focal actor characteristics (both network and non-network) are uncorrelated. Randomly pairing individuals helps to address the problem of selection by breaking the correlation between peer and focal individual characteristics. To address reflection, the researcher requires longitudinal network data for both the focal individual and her peers. Longitudinal data allow us to construct lagged measures of peer networks. Lagged measures of peer networks are less susceptible to the problem of simultaneity or unobserved environmental causes.

Using data that meet these two requirements, we test our claim that peers affect network growth and position. Our data are derived from a natural experiment at an engineering college in India (Hasan and Bagde 2013). Students at this college were randomly assigned to roommates in their first and second years. We use the second-year random assignment of 2,113 students to roommates combined with three years of social network data to examine whether peers—randomly assigned roommates in this case—cause a focal student's network to grow. Specifically, we test whether having a roommate with a large preexisting network causes the focal student's network to subsequently grow. We find evidence that roommates with large networks do indeed cause the focal student to experience a significant creation of new ties. The focal student is found to create new ties not only to members of her roommates' networks but also to individuals not previously connected to her roommate. Finally, our results indicate that well-connected roommates also increase the likelihood that the focal student serves as a bridge in the network and moves closer to the center of the larger network structure.

Research Design

To test whether peers affect network growth, we analyze data on the networks of one cohort of students at an engineering college in a southern state of India. The data were collected over a three-year period, from the students' first academic year until their third academic year. The population of students consists of the college's first entering cohort. In that year, the college enrolled 2,161 students. Of these students, 2,122 were still enrolled in the third academic year. The final sample consists of a comprehensive set of variables describing the background characteristics and academic performance of 2,113 students, accounting for 97.7% of originally enrolled students and 99.5% of all enrolled students in the third year. Besides data on social background and academic performance, college administrators provided us with data on the dormitory assignments of students, which were, as per standard university policy, exogenous and random.

In addition to the academic and demographic data, we surveyed students each year about their social networks. We conducted the first social network survey during the third week of the first academic year. At this point, the college had already assigned students to their first-year dormitory rooms and students had begun taking classes, although no examinations were yet conducted. We distributed social network surveys to all enrolled students. The survey asked students about their friendship networks ("list the first name, last name, hostel, and district of the students at [the college] who you consider your close friends") and their "study-partner" networks ("list the first name, last name, hostel, and district of the students with whom you study"). We note that the study-partner networks are *informal* study partners and that in many cases study-partner relations are asymmetric. One student may list another student as a study partner, but that student may not necessarily list the first as a study partner. Thus, in many respects, the study-partner relation is akin to informal task-networks and instrumental relations in other organizational contexts. Students could list up to 12 individuals for each relation and were asked to list as few or as many names as they would like. Any student who wanted to participate could return the completed survey to the researcher by the end of the day. In the second and third years, the students were requested to complete the same survey in the same manner. In both years the surveys were also conducted during the third week of each academic year. We achieved relatively high levels of response for all three years. We achieved effective response rates of 86% in the first year, 73% in the second year, and 90.4% in the third year.[1] The survey response rates compare quite favorably to other network surveys. Moreover, our sample sizes are significantly larger than most network surveys, either cross-sectional or longitudinal. We estimated logistic regressions to predict response in the first and third years, the years that we used to construct our independent and dependent variables. We found no consistent relationship between survey response and student characteristics. We present summary statistics for the main variables used in our analyses in Table 1.

A key feature of our data is the university policy to randomly assign students to dormitory rooms in the first and second academic years. Unlike American universities, where students can state preferences for roommates with certain characteristics (e.g., smokers versus nonsmokers), this Indian university does not conduct such preference surveys. However, we note that the random assignment was conditional on gender: the dormitories were not coeducational, and male and female students were never assigned to the same dormitory room. Our results are therefore all within-gender estimates. A university administrator conducted the random assignment using Microsoft Excel in the following manner: first, the administrator partitioned rooms into those for male and female students; second, the administrator randomly assigned

[1] We hired a company in India to hand-transcribe each of the surveys into electronic form. After surveys were transcribed, each name in each survey was matched to a name in the electronic database of students, first algorithmically and then manually. Across the three years of the survey we were able to match over 93% of the names in the social network survey responses to names in the administrative database. The matching was aided by the fact that 2,156 of 2,160 names in the administrative database were unique. We were able to disambiguate the remaining four names using data on the individual's district of origin and dormitory room name.

Table 1 Summary Statistics of Key Dependent and Independent Variables Used in the Analysis

Variable	Mean	SD	Min	Max
Female	0.442	0.497	0.000	1.000
SC	0.186	0.390	0.000	1.000
ST	0.080	0.271	0.000	1.000
BC-A	0.134	0.341	0.000	1.000
BC-B	0.252	0.434	0.000	1.000
BC-C	0.011	0.106	0.000	1.000
BC-D	0.176	0.381	0.000	1.000
Muslim	0.040	0.197	0.000	1.000
Open Category	0.120	0.325	0.000	1.000
Sem 2 (*Self*)	0.006	0.997	−2.376	2.929
Own HS Score	0.002	1.001	−5.723	2.168
Roommates' Study Net Size	5.459	2.223	0.500	19.000
Roommates' Friend Net Size	6.791	2.356	0.500	26.500
Own Friends (*Year 1*)	6.791	4.764	0.000	34.000
Own Friends (*Year 2*)	6.159	3.794	0.000	26.000
Own Friends (*Year 3*)	7.550	4.201	0.000	32.000
Own Study Net (*Year 1*)	5.459	4.547	0.000	32.000
Own Study Net (*Year 2*)	5.088	3.879	0.000	27.000
Own Study Net (*Year 3*)	5.690	3.364	0.000	26.000
Number of Roommates	6.276	4.020	2.000	28.000
Roommates' Academic Performance	0.006	0.466	−1.341	1.959
In Roommates' Network	0.155	0.447	0.000	5.000
Out of Roommates' Network	4.995	3.030	0.000	23.000

Note. $N = 2{,}113$.

students to rooms until he had assigned all students to a room. In the first academic year, the college had 92 dormitory rooms with approximately 24 students assigned to a room. In the second academic year, the college had 394 dormitory rooms with a median of 6 students assigned to a room.

Below we describe the construction of our key dependent and independent variables.

Dependent Variables: Measures of Third-Year Network Structure

Own Study and Friendship Network Size. Our primary dependent variable is the size of a focal student's network in her third year at the college. To measure network size, we use *indegree* centrality. We calculate the indegree by counting the number of other individuals in the network who consider the focal student a friend or a study partner. The indegree is a useful and theoretically relevant measure of centrality for several reasons. First, researchers have linked the indegree measure to a variety of outcomes for individuals as well as organizations. Authors have found that network size correlates to outcomes such as job satisfaction (Roberts and O'Reilly 1979) and power. Second, research finds that the indegree correlates with, and is a component of, other measures of centrality such as closeness and betweenness, which capture the size of an individual's network as well as the extent to which an actor is linked to disparate parts of the broader network structure. Research has found that indegree centrality has an average correlation of 0.62 with betweenness centrality and 0.55 with closeness centrality (Valente et al. 2008). Third, the indegree has several properties that make it a relatively robust measure of centrality, even in the presence of incomplete or imperfect survey response. In our specific case, although we have data on a relatively large, and for the most part complete, population of individuals, we construct our indegree measures with missing data. Our choice of indegree alleviates worries arising from such issues. Research finds that the indegree is robust to sampling (Costenbader and Valente 2003); even when 50% of data is missing, the sampled indegree has a 0.9 correlation to indegree constructed with complete data. Even when only 10% of the network data is available, the sampled indegree has a correlation of 0.5 to an indegree computed with the complete network. Since our response rates for the first and third years are 86% and 91%, respectively, we are confident that our network variables accurately represent each student's proportional network size. We calculated the indegree measure variable using both the friendship and study-partner network for the 2,113 students in our sample. These are called *Own Friends* and *Own Study Net*, respectively, for each of the three years in our analysis. Although the friendship and study-partner indegree correlate (corr $= 0.54$, $p < 0.01$), they do not completely overlap. Thus, we estimate separate models with each measure as our dependent variable.[2]

In addition to the main dependent variables described above, we also constructed two additional measures of indegree. The first measure is the count of a focal student's incoming third-year study partners meeting two conditions: (1) they were previously study partners or friends of her roommates, but (2) they were previously not study partners or friends of her own. We call this variable *In Roommates' Network*. We use this variable to measure the number of ties created in the third year. The second measure, *Out of Roommates' Network*, is the count of a focal student's incoming third-year study partners who meet the following conditions: (1) they were previously not in the network of a focal student, and (2) they were previously not in her roommates' network. We use this

[2] We present histograms of indegrees in the study-partner and friendship networks in the online appendix (available as supplemental material at http://dx.doi.org/10.1287/mnsc.2014.2109) as Figures A1 and A2, respectively. The figures suggest that the dependent variables follow count distributions such as Poisson or negative binomial. Since we are using network data, we have chosen to use negative binomial models, allowing us to account for overdispersion. As a check on the robustness of our results, we also conducted all our analyses using Poisson models and linear regressions. Our coefficient estimates were equivalent in both magnitude and level of significance under these specifications.

variable to model whether well-connected roommates increased the network size of focal students beyond the closure effects.

Farness Centrality in the Third Year. To measure an individual's distance to other actors in the network, we compute farness centrality, the inverse of the more commonly used *closeness* centrality measure. We calculate farness by determining the shortest path between each pair of actors i and j in graph g, which counts the minimum number of edges separating the two actors. Once we calculate shortest paths, denoted by $dist(i, j)$, we can compute a farness measure for each actor by averaging across all ij pairs for each actor i, where n is the number of nodes in the network. Higher values of farness indicate that the focal actor is more distant, on average, from all others in the network than someone with a lower score. As expected, farness centrality and indegree in the study-partner network are negatively correlated at $\rho = -0.469$ and with statistical significance at the $p < 0.001$ level. We present the formula for farness in Equation (1):

$$Farness_i(g) = \sum_j \frac{dist(i, j)}{n-1}.$$

Betweenness Centrality in the Third Year. We computed betweenness centrality to measure how often an individual serves as a bridge within the larger network. Betweenness measures how frequently a focal actor lies on the shortest path between all other actors in the network. If, for instance, four shortest paths link two actors j and k, and actor i lies on this shortest path all four times, then i is a critical bridge. Without i, information transfer between j and k is difficult. On the other hand, if i never lies on the shortest path, then i is less critical in the channeling of information between the two parties j and k. We compute betweenness by summing across the proportion of instances that i serves as a bridge between all jk pairs, denoted by $\sigma_{jk}(i)/\sigma_{jk}$. We find that betweenness and indegree in the third-year study-partner network have a correlation of $\rho = 0.475$. We present the formula for betweenness in Equation (2):

$$Betweenness_i(g) = \sum_{jk} \frac{\sigma_{jk}(i)}{\sigma_{jk}}.$$

Independent Variables: Peers' Network Structure

We constructed two independent variables for our analyses. To construct these variables, we define a student's relevant peers as her randomly assigned second-year roommates. We use these roommates' indegree in their friendship and study-partner networks as measured in the first academic year (before roommate assignment) to construct our peer network variables. For each student we created two variables: *Roommates' Study Net Size* and *Roommates' Friend Net Size*. The two variables measure the average indegree of the second-year roommates' study-partner and friendship network indegree in their first year, respectively. Two research design considerations motivate the choice of second-year roommates versus first-year roommates. First, we measure indegree in the first year before the treatment (random assignment of roommates for the second year) occurs, whereas we measure second-year indegree after the treatment. Thus, the use of second-year indegree raises problems of reflection, making our causal estimates less credible. Moreover, we use the second-year roommates as the relevant peer group because we do not have pretreatment network structure measures for first-year roommates. Although this gives a more conservative test, we can make causal claims. Statistical tests evaluating the effectiveness of the random assignment of roommates can be found in Tables A1 and A2 in the online appendix. In our robustness checks, we also conducted less conservative tests using roommates' networks in the second year as well; these results can be found in Table A3 in the online appendix.

Control Variables

In addition to the two key independent variables mentioned above, we wanted to examine whether other characteristics of second-year roommates affected a focal student's network structure. To do this, we create a variable measuring second-year roommates' average performance on the high school board exam called *Roommates' Academic Performance*. We use this variable to examine whether alternative quality-based mechanisms might mediate the relationship between roommates' networks and a focal student's network.

The strongest feature of our analysis is random assignment. However, statisticians suggest that researchers should still include control variables when estimating the relationship between a treatment variable and the dependent variables (Imai et al. 2008, Gelman 2011). We can control for several pretreatment characteristics of the focal student because our data include detailed background information about students, their academic performance in Semester 2 (*Sem 2 Self*) and in high school (*Own HS Score*), as well as their networks. The main estimations include variables indicating whether a student is female and their caste category,[3] as well as the focal student's prior network structure and academic performance. We account for prior network structure by including the focal student's first- and second-year indegree in the two types of networks we measured.

[3] In our models we include seven fixed effects for caste category: scheduled castes (*SC*), scheduled tribes (*ST*), four backward caste categories (*BC-A*, *BC-B*, *BC-C*, *BC-D*), and Muslims (*Muslim*). The omitted category is *Open Category*, representing what are considered upper or forward castes.

Because academic performance may determine network structure—with high-achieving students having more study partners or friends—our models include controls for a focal student's prior academic performance on the standardized high school board exam, first-year grade point average (GPA), and second-year GPA. The measures of academic achievement are comparable across students. In each case, the college evaluated students using the same examination in subjects including mathematics, physics, chemistry, English, and the regional language. We also include the number of second-year roommates for each student to control for the simple effect of having many roommates. Finally, our models include fixed effects for each student's district of origin. All background data are derived from archival records provided by the university. We have clustered our standard errors at the second-year dormitory room level to account for multiple observations within a room.

Results

Do Peers Cause Network Growth?

We begin our empirical analysis by examining the central claim of this paper: peers cause network growth. We argued that focal students with well-connected roommates would experience more network growth than those who without. We test this claim using two types of network ties: informal studying relations and friendship relations. In our empirical test, we estimate a series of negative binomial models that regress the indegree of the focal student on the *Roommates' Study Net Size* and *Roommates' Friend Net Size* variables. All models include a substantial set of controls including those for the focal student's caste, prior academic performance, gender, and district of origin.

We present the first set of models in Table 2, which examines growth in the study-partner network. In column (1), we find that having roommates with large networks does increase the size of the focal student's third-year study network. The key result holds even when we control for the number of roommates a student has. In columns (2)–(5), we separately include control variables for the focal student's study-partner and friendship network size in the student's first and second years. In each case, we see a significant and positive relationship between these variables and the dependent variable. As expected, students who previously had large networks continue to have them. Furthermore, the magnitude of the *Own Study Net* (*Year 2*) variable is larger than that of the *Own Study Net* (*Year 1*) variable. In column (6), we see that *Roommates' Study Net Size* remains positive and statistically significant even when we include the entire set of own network size controls. What is more striking is that the magnitude of the *Roommates' Study Net Size* variable remains stable across all six models. The magnitude of this main effect is not insubstantial. The magnitude of one unit change in roommates' network size is approximately 48% (0.013/0.029) of the increase in own indegree in the first year and 29% (0.013/0.045) of the increase in own indegree in the second year. We argue that this test provides strong evidence that differences in network structure can be

Table 2 Negative Binomial Models of Own Study-Partner Network Size in Year 3 on Average Size of Roommates' Study-Partner Network in Year 1

	(1)	(2)	(3)	(4)	(5)	(6)
Roommates' Study Net Size	0.014**	0.014**	0.013**	0.013**	0.013**	0.013**
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)
Number of Roommates	−0.007	−0.005	−0.006	−0.005	−0.005	−0.004
	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)
Own Study Net (*Year 1*)		0.029**				0.012**
		(0.003)				(0.004)
Own Study Net (*Year 2*)			0.045**			0.028**
			(0.003)			(0.003)
Own Friends (*Year 1*)				0.027**		0.006
				(0.003)		(0.004)
Own Friends (*Year 2*)					0.047**	0.026**
					(0.003)	(0.003)
Constant	1.628**	1.428**	1.392**	1.431**	1.348**	1.197**
	(0.067)	(0.071)	(0.068)	(0.072)	(0.067)	(0.070)
Background controls	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
ln alpha (constant)	−2.038**	−2.204**	−2.327**	−2.198**	−2.323**	−2.540**
	(0.085)	(0.088)	(0.102)	(0.092)	(0.098)	(0.113)

Notes. Standard errors are clustered at the dormitory room level. $N = 2{,}113$.
$^{*}p < 0.05$; $^{**}p < 0.01$ (all tests are two-tailed).

Table 3 **Negative Binomial Models of Own Friendship Network Size in Year 3 on the Average Size of Roommates' Friendship Network in Year 1**

	(1)	(2)	(3)	(4)	(5)	(6)
Roommates' Friend Net Size	0.009*	0.009*	0.008*	0.008*	0.007*	0.007*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)
Number of Roommates	−0.005	−0.003	−0.004	−0.003	−0.002	−0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Own Study Net (Year 1)		0.029**				0.007*
		(0.003)				(0.003)
Own Study Net (Year 2)			0.041**			0.013**
			(0.003)			(0.003)
Own Friends (Year 1)				0.030**		0.009*
				(0.003)		(0.003)
Own Friends (Year 2)					0.066**	0.052**
					(0.003)	(0.003)
Constant	1.966**	1.770**	1.766**	1.745**	1.579**	1.480**
	(0.061)	(0.062)	(0.055)	(0.062)	(0.056)	(0.056)
Background controls	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
ln alpha (constant)	−1.976**	−2.130**	−2.183**	−2.164**	−2.543**	−2.649**
	(0.067)	(0.072)	(0.079)	(0.074)	(0.102)	(0.108)

Notes. Standard errors are clustered at the dormitory room level. $N = 2{,}113$.
$^{*}p < 0.05$; $^{**}p < 0.01$ (all tests are two-tailed).

caused by exogenous peer characteristics. The lagged measures of roommate network structure and the random assignment reduce the key inferential concerns arising from the reflection and selection problems (Manski 1999).[4]

In Table 3 we estimate a similar series of models for growth in the friendship network. Here, we regress the focal student's friendship indegree in the third year on the roommates' network size. We similarly find that roommates' friendship network size significantly predicts the focal student's own friendship indegree. These results, similar to those in Table 2, are robust in both magnitude and significance even after controlling for a student's number of roommates, her own network size in years 1 and 2, and a suite of individual control variables. However, we see that the friendship indegree of roommates has a smaller effect on friendship network growth than within the study-partner network. The main effect is 30% the size of the focal student's own first-year network and 10% the size of second-year friendship network.

Qualitatively, we observe differences in the magnitude of effects across the friendship and study-partner networks. Several potential explanations for such a difference exist. First, it may be the case that instrumental network ties such as studying relations are easier to create through the mechanisms we proposed than is friendship. Friendships may require significantly more similarity between actors than the random assignment of students to dormitory rooms induces (Lincoln and Miller 1979). A second explanation is more structural. We observe differential yearly churn across network types that may affect whether students form new network ties based on their roommates' networks. We find, for instance, that across the first and second years, 41% of friendship relations persist, whereas 17.3% of study relations do. To be sure, there is considerable variability: some students are able to maintain more of their connections than others. What is also striking is that, despite high churn, the size of the networks across years is quite stable. The mean change in network size across the first and third years is approximately 0.75 for the friendship network and 0.22 for the study-partner network. This pattern of results suggests that individuals drop as well as add network ties. New ties appear to replace ones that actors drop; thus, new ties allow network size to remain constant over time.

[4] We reanalyzed the effect of roommates' networks on the focal student's own third-year networks using roommates' second-year networks (compared with first-year networks). Although this approach contaminates the causal inference because of the reflection problem—since it becomes difficult to ascertain whether the focal student's network affected her roommates, or vice versa—it provides an additional robustness check on the validity of our results. We find that roommates' second-year networks have a significant effect on the focal student's third-year networks, and the magnitude is larger than that of roommates' first-year networks, although not significantly so. The beta coefficient for the first-year effect is 0.013, versus 0.016 for the second-year coefficient, with the standard errors being 0.004 for both estimates. Thus, we have a z-statistic of approximately 0.53. Thus, we cannot reject the null hypothesis that the coefficients are of the same magnitude.

Well-connected roommates are an important source of connections and resources for this process of network renewal. More specifically, it appears that the high churn in the study-partner network produces more structural "space" compared with the friendship network. A student is able to fill this space produced by the churn in the study-partner network more readily when her roommate has many connections to whom the focal student can be introduced. A further pattern in our data—that individuals with lower churn have larger networks, with a correlation of $\rho = 0.22$ and statistical significance at the $p < 0.001$ level—appears to give further support to the idea that well-connected roommates are better able to help the focal student renew her network. Thus, we conjecture that one reason that differential effect sizes across network types exist is because of the varying rates of churn across them.

Finally, although our measures of roommates' network size are credibly exogenous, they may still reflect different, although correlated, characteristics of roommates. To examine whether we could explain our effects by using other roommate characteristics, we estimated additional models in Table 4. One competing explanation is that the high indegree of roommates simply reflects high levels of prior academic performance. Thus, after controlling for roommates' prior performance, roommates' network size should no longer matter. To test this claim, we include lagged measures of roommates' average performance in high school and college. We present these results in columns (1) and (2), respectively. The coefficients for these variables are not significant, nor do they substantively change our main results. The *Roommates' Study Net Size* variables are not significantly different from each other, nor are they different from the coefficient in Table 1, column (6). In column (4) of Table 4, we examine whether the increase in study-partner network indegree is a result of having roommates with large study-partner networks *or* just large networks. We see that roommates' friendship indegree has a statistically significant effect on the size of a student's study-partner network, and the magnitude of this effect is similar to the effects seen in columns (1)–(3). Finally, we examined whether a decreasing marginal effect of roommates' indegree on growth exists. To test this possibility, we included the square of *Roommates' Study Net Size*. The results appear to suggest that, within the range of our data, the main effect is primarily linear.

Table 4 Negative Binomial Models of Examining Alternative Mechanisms for the Relationship Between Own Study-Partner Network Size in Year 3 on Average Size of Roommates' Study-Partner Network in Year 1

	(1)	(2)	(3)	(4)
Roommates' Study Net Size	0.013** (0.004)	0.012* (0.005)		0.012** (0.005)
(*Roommates' Study Net*)^2				−0.000 (0.000)
Own Study Net (*Year 1*)	0.012** (0.004)	0.012** (0.004)	0.012** (0.004)	0.012** (0.004)
Own Study Net (*Year 2*)	0.028** (0.003)	0.028** (0.003)	0.028** (0.003)	0.029** (0.003)
Own Friends (*Year 1*)	0.006 (0.004)	0.006 (0.004)	0.006 (0.004)	0.006 (0.004)
Own Friends (*Year 2*)	0.026** (0.003)	0.026** (0.003)	0.026** (0.003)	0.026** (0.003)
Number of Roommates	−0.004 (0.003)	−0.004 (0.003)		−0.004 (0.003)
Roommates' Prior Academic Performance (HS)	−0.017 (0.012)			
Roommates' Prior Academic Performance (College)		−0.006 (0.022)		
Roommates' Friend Net Size			0.011** (0.004)	
Background controls	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes
Constant	1.194** (0.070)	1.203** (0.071)	1.166** (0.069)	1.207** (0.071)
ln alpha (constant)	−2.544** (0.114)	−2.540** (0.113)	−2.532** (0.113)	−2.540** (0.113)

Notes. Standard errors are clustered at the dormitory room level. HS, high school. $N = 2{,}113$.
$^*p < 0.05$; $^{**}p < 0.01$ (all tests are two-tailed).

Empirical Extensions

Is There Growth Beyond the Roommate's Network? The previous estimates provide causal evidence that well-connected roommates cause a focal student's network to grow. However, peers could cause network growth in several ways. Although identifying precise mechanisms is something we cannot do with our data, we can examine whether peer networks cause the creation of ties to individuals within and outside a peer's network. To examine whether interactions with peers cause "closure growth" as a result of introductions to people in a roommate's preexisting network, we construct a count of a focal student's third-year study partners that meet two conditions: (1) they were previously study partners of their roommates, but (2) they were not previously study partners of her own. We call this variable *In Roommates' Network*. We estimate negative binomial models regression of this variable on the roommates' study-partner network size. In column (1) of Table 5, we include only the size of the roommates' study-partner network and the background controls. This variable's effect in our model is positive and significant. In column (2), we include the number of roommates a student has as well as her own prior network size in our model. This variable is statistically significant and suggests that having more roommates increase one's

Table 5 Negative Binomial Models of Own Study-Partner Network Closure Growth in Year 3 on Average Size of Roommates' Study-Partner Network in Year 1

	(1)	(2)	(3)
Roommates' Study Net Size	0.110**	0.152**	
	(0.037)	(0.035)	
Own Study Net (*Year 1*)		0.010	0.005
		(0.024)	(0.023)
Own Study Net (*Year 2*)		0.016	0.012
		(0.016)	(0.016)
Own Friends (*Year 1*)		0.028	0.030
		(0.020)	(0.020)
Own Friends (*Year 2*)		0.038*	0.035
		(0.018)	(0.018)
Number of Roommates		0.095**	0.098**
		(0.012)	(0.012)
Roommates' Friend Net Size			0.154**
			(0.022)
Constant	−2.171**	−3.745**	−4.004**
	(0.372)	(0.406)	(0.388)
Background controls	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes
ln alpha (constant)	0.153	−0.718	−0.706
	(0.459)	(0.489)	(0.459)

Notes. Standard errors are clustered at the dormitory room level. $N = 2{,}113$.
$^{*}p < 0.05$; $^{**}p < 0.01$ (all tests are two-tailed).

potential pool of new connections and thus network size. Interestingly, apart from a small effect of friendship indegree in the second year, all coefficients are small and insignificant. This suggests that peer networks are the primary cause of closure-induced tie creation in the study-partner network. The pattern of results indicates strong support for closure-induced tie creation triggered by the administrative intervention of assigning roommates. Finally, in column (3), we examine whether this result generalizes to having friends with large friendship networks as well. We find that the results hold; having roommates with many friends also increases one's chances of creating study-partner ties through introductions and then closure. We also estimated dyadic models examining the probability of a student i forming a tie to another student, j, in the third year if j is a study partner of i's roommate. We find that the odds of such a tie forming are 2.39 higher than they would be if i and j were not connected via the roommate.

Peers can also help the focal actor to create ties to individuals not already in that peer's network. Such growth does not happen because of closure or introductions but because of other factors that could include learning networking skill, introduction to a new foci or setting, or the mere association with someone popular. Although the indirect mechanisms may also increase the size of the *In Roommates' Network* variable, growth in the number of connections to individuals *not* previously connected to the focal student's roommate would constitute stronger evidence. To test for indirect growth, we regress a variable called *Out of Roommates' Network* on *Roommates' Study Net Size*. We present estimates from this model in column (1) of Table 6. We indeed find some evidence that individuals who have well-connected roommates experience growth in their networks that is not directly related to the introductions that the roommates provide. In column (2), we include controls for own network size in prior years; as expected, we find that the presence of new ties relates to the size of one's own prior network. In column (3), we include the *In Roommates' Network* variable in our model to test whether growth occurs because of introductions made by roommates' connections. However, because the *In Roommates' Network* variable overlaps with *Roommates' Study Net Size*, we cannot precisely determine whether the learning, associational mechanism, or a chain of introductions is the cause of the growth resulting from the *Out of Roommates' Network* variable. In column (4), we estimate this model using *Roommates' Friend Net Size*. Although this variable is positive, it is not statistically significant. Finally, in column (5), we include fixed effects for the classrooms to which the focal student belonged in prior semesters to model differential opportunities for interaction. We find that the main effect persists even with these controls. Together, our results provide evidence for both closure and nonclosure network growth.

Do Peer Networks Affect Structural Position?

Next, we relate *Roommates' Study Net Size* to the focal student's distance to others in the larger network (*Farness*) and the extent to which she serves as a bridge (*Betweenness*). We present the results of these analyses in Table 7. In column (1), we estimate a model where we regress the network farness of the focal student on the *Roommates' Study Net Size* variable. The effect for size is negative and statistically significant, suggesting that students whose roommates have large networks are more likely to connect them to distant parts of the network, thereby bringing them closer to all others. We also control for the distance between the focal student and the peers in the first-year network; however, this distance effect is not statistically significant. In column (2), we examine whether roommates' network size affects a student's *Betweenness*—i.e., how often a student serves as a bridge in the network. We find that the betweenness of a student increases significantly if she has a well-connected roommate. Together, these results indicate that growth in a focal student's network also reconfigures her network by pulling her more to the center of the network and giving her more opportunities to serve as a network bridge.

Table 6 Negative Binomial Models of Own Study-Partner Network Nonclosure Growth in Year 3 on Average Size of Roommates' Study-Partner Network in Year 1

	(1)	(2)	(3)	(4)	(5)
Roommates' Study Net Size	0.012**	0.011**	0.008*		0.010*
	(0.004)	(0.003)	(0.004)		(0.005)
Own Study Net (Year 1)		0.001	0.000	0.001	0.004
		(0.003)	(0.003)	(0.003)	(0.004)
Own Study Net (Year 2)		0.031**	0.030**	0.030**	0.032**
		(0.003)	(0.003)	(0.003)	(0.003)
Own Friends (Year 1)		0.002	0.003	0.003	0.004
		(0.003)	(0.003)	(0.003)	(0.004)
Own Friends (Year 2)		0.026**	0.024**	0.025**	0.023**
		(0.003)	(0.003)	(0.003)	(0.004)
Number of Roommates		−0.011**	−0.013**	−0.011**	−0.013**
		(0.003)	(0.003)	(0.003)	(0.003)
In Roommates' Network			0.058**		
			(0.022)		
Roommates' Friend Net Size				0.006	
				(0.003)	
Constant	1.424**	1.206**	1.196**	1.228**	−0.704***
	(0.046)	(0.044)	(0.059)	(0.048)	(0.217)
Background controls	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes
Classroom fixed effects	No	No	No	No	Yes

Notes. Standard errors are clustered at the dormitory room level. $N = 2{,}113$.
$^{*}p < 0.05$; $^{**}p < 0.01$ (all tests are two-tailed).

Table 7 Linear Regression Models of Own Farness and Betweenness Centrality in Year 3 on Roommate Distance and Average Size of Roommates' Study-Partner Network in Year 1

	(1) *Farness*	(2) *Betweenness*
Own Study Net (Year 1)	−0.003	66.361
	(0.003)	(61.212)
Own Study Net (Year 2)	−0.015**	313.068**
	(0.002)	(65.725)
Own Friends (Year 1)	−0.004	121.192*
	(0.003)	(53.822)
Own Friends (Year 2)	−0.011**	122.436*
	(0.003)	(55.508)
Number of Roommates	0.004	−0.377
	(0.004)	(28.152)
Roommates' Study Net Size	−0.018**	70.609**
	(0.005)	(26.288)
Distance Between Focal Student and Peer	−0.028	
	(0.021)	
Constant	4.451**	2,368.776**
	(0.121)	(682.094)
Background controls	Yes	Yes
District fixed effects	Yes	Yes
Adjusted R^2	0.088	0.072

Notes. Standard errors are clustered at the dormitory room level. $N = 2{,}113$.
$^{*}p < 0.05$; $^{**}p < 0.01$ (all tests are two-tailed).

Discussion and Conclusion

Social network theory treats processes related to how networks emerge and how they affect the outcomes of individuals and organizations (Borgatti and Halgin 2011). In more recent years, there has been much attention in the scholarly literature on developing empirical tests of the causal effects of social networks on economic and organizational outcomes (Mouw 2006). However, to our knowledge, there has been much less research examining the causal mechanisms that drive network formation (for exceptions, see Marmaros and Sacerdote 2006 and Kleinbaum 2012). In this article, we take advantage of the random assignment of students to dormitories in an Indian college and data on their social networks over three consecutive years to provide more credible evidence of the causes of network formation. The analysis in this article formalizes and examines one source of network change: the effect of peers on network growth. Our results provide evidence for network theory's central assumption that network position cannot simply be reduced to individual attributes, dispositions, and strategies. First, we find that interacting with well-connected roommates leads to the growth of a student's social network. Additional analysis suggests that a student's network primarily grows because her roommates introduce her to members of their preexisting networks, helping her renew her network after she loses connections as a result of churn. Finally, we find evidence that interacting with well-connected roommates also increases the extent to which a student becomes a network bridge and moves closer to the network's center.

In concluding, we would like to note several important limitations of the analyses presented in this article. First, our data—and, by extension, our analyses—derive from a single college in India. Although we believe our results are internally valid, problems of generalizability do exist. The setting limits our ability to generalize to other types of organizations and to other cultural contexts outside India. We are less worried about the college setting. Colleges are important venues for individuals to meet new people and expand their networks. Many students choose colleges and graduate schools for the explicit goal of expanding their networks. The choice of MBA programs and even undergraduate universities often follows this logic. Nevertheless, we believe we have outlined processes general enough to apply to a wide variety of settings and provided a novel and generalizable research design for studying network dynamics. Moving forward, we think there is much promise in using the tools of field and natural experiments to understand network formation and dynamics. In particular, we think that multiple sequential random assignments, such as the ones that we used in this study, are a fruitful way to understand the impact of altering networks on both personal and macronetwork evolution. Finally, we think that the exogenous network change caused by multiple random assignments may be a useful instrument for understanding the causal effect of network structure on advantage in organizations and markets.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mnsc.2014.2109.

Acknowledgments

The authors thank members of the macro-organizational behavior group at the Stanford Graduate School of Business; seminar participants at University of Chicago's Booth School of Business, MIT Sloan School of Management, and Stanford's Political Methodology Workshop; and the director and staff of the engineering college studied here for their assistance and suggestions. The authors thank Carnegie Mellon University's Heinz College and Stanford University's Graduate School of Business for their generous support.

References

Aral S, Walker D (2011) Creating social contagion through viral product design: A randomized trial of peer influence in networks. *Management Sci.* 57(9):1623–1639.

Borgatti SP, Halgin DS (2011) On network theory. *Organ. Sci.* 22(5):1168–1181.

Borgatti SP, Mehra A, Brass DJ, Labianca G (2009) Network analysis in the social sciences. *Science* 323(5916):892–895.

Brass DJ (1984) Being in the right place: A structural analysis of individual influence in an organization. *Admin. Sci. Quart.* 29(4):518–539.

Burt RS (2004) Structural holes and good ideas. *Amer. J. Sociol.* 110(2):349–399.

Burt RS (2009) *Structural Holes: The Social Structure of Competition* (Harvard University Press, Cambridge, MA).

Burt RS (2012) Network-related personality and the agency question: Multirole evidence from a virtual world. *Amer. J. Sociol.* 118(3):543–591.

Burt RS, Ronchi D (2007) Teaching executives to see social capital: Results from a field experiment. *Soc. Sci. Res.* 36(3):1156–1183.

Carrell SE, Sacerdote BI, West JE (2013) From natural variation to optimal policy? The importance of endogenous peer group formation. *Econometrica* 81(3):855–882.

Costenbader E, Valente TW (2003) The stability of centrality measures when networks are sampled. *Soc. Networks* 25(4):283–307.

Currarini S, Jackson MO, Pin P (2009) An economic model of friendship: Homophily, minorities, and segregation. *Econometrica* 77(4):1003–1045.

Feld SL (1981) The focused organization of social ties. *Amer. J. Sociol.* 86(5):1015–1035.

Feld SL (1982) Social structural determinants of similarity among associates. *Amer. Sociol. Rev.* 47(6):797–801.

Fernandez RM, Weinberg N (1997) Sifting and sorting: Personal contacts and hiring in a retail bank. *Amer. Sociol. Rev.* 62(6): 883–902.

Gelman A (2011) Causality and statistical learning. *Amer. J. Sociol.* 117(3):955–966.

Gould RV (2002) The origins of status hierarchies: A formal theory and empirical test. *Amer. J. Sociol.* 107(5):1143–1178.

Granovetter MS (1973) The strength of weak ties. *Amer. J. Sociol.* 78(6):1360–1380.

Hartmann WR, Manchanda P, Nair H, Bothner M, Dodds P, Godes D, Hosanagar K, Tucker C (2008) Modeling social interactions: Identification, empirical methods and policy implications. *Marketing Lett.* 19(3–4):287–304.

Hasan S, Bagde S (2013) The mechanics of social capital and academic performance in an Indian college. *Amer. Sociol. Rev.* 78(6):1009–1032.

Ibarra H (1992) Homophily and differential returns: Sex differences in network structure and access in an advertising firm. *Admin. Sci. Quart.* 37(3):422–447.

Imai K, King G, Stuart EA (2008) Misunderstandings between experimentalists and observationalists about causal inference. *J. Roy. Statist. Soc. Ser. A* 171(2):481–502.

Kleinbaum AM (2012) Organizational misfits and the origins of brokerage in intrafirm networks. *Admin. Sci. Quart.* 57(3):407–452.

Kossinets G, Watts DJ (2006) Empirical analysis of an evolving social network. *Science* 311(5757):88–90.

Kossinets G, Watts DJ (2009) Origins of homophily in an evolving social network. *Amer. J. Sociol.* 115(2):405–450.

Lerner J, Malmendier U (2013) With a little help from my (random) friends: Success and failure in post-business school entrepreneurship. *Rev. Financial Stud.* 26(10):2411–2452.

Lincoln JR, Miller J (1979) Work and friendship ties in organizations: A comparative analysis of relation networks. *Admin. Sci. Quart.* 24(2):181–199.

Manski CF (1993) Identification of endogenous social effects: The reflection problem. *Rev. Econom. Stud.* 60(3):531–542.

Manski CF (1999) Analysis of choice expectations in incomplete scenarios. *J. Risk Uncertainty* 19(1–3):49–66.

Marmaros D, Sacerdote B (2006) How do friendships form? *Quart. J. Econom.* 121(1):79–119.

McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a feather: Homophily in social networks. *Annual Rev. Sociol.* 27:415–444.

Mehra A, Kilduff M, Brass DJ (2001) The social networks of high and low self-monitors: Implications for workplace performance. *Admin. Sci. Quart.* 46(1):121–146.

Mouw T (2006) Estimating the causal effect of social capital: A review of recent research. *Annual Rev. Sociol.* 32:79–102.

Nanda R, Sørensen JB (2010) Workplace peers and entrepreneurship. *Management Sci.* 56(7):1116–1126.

Obstfeld D (2005) Social networks, the *tertius iungens* orientation, and involvement in innovation. *Admin. Sci. Quart.* 50(1): 100–130.

Podolny JM (2001) Networks as the pipes and prisms of the market. *Amer. J. Sociol.* 107(1):33–60.

Podolny JM, Baron JN (1997) Resources and relationships: Social networks and mobility in the workplace. *Amer. Sociol. Rev.* 62(5):673–693.

Reagans RE, Zuckerman E, McEvily B (2007) On firmer ground: The collaborative team as strategic research site for verifying network-based social capital hypotheses. Rauch JE, ed. *The Missing Links: Formation and Decay of Economic Networks*, Chap. 6 (Russell Sage Foundation, New York), 147–182.

Roberts KH, O'Reilly CA III (1979) Some correlations of communication roles in organizations. *Acad. Management J.* 22(1): 42–57.

Sacerdote B (2001) Peer effects with random assignment: Results for Dartmouth roommates. *Quart. J. Econom.* 116(2):681–704.

Sasovova Z, Mehra A, Borgatti SP, Schippers MC (2010) Network churn: The effects of self-monitoring personality on brokerage dynamics. *Admin. Sci. Quart.* 55(4):639–670.

Small ML (2010) *Unanticipated Gains: Origins of Network Inequality in Everyday Life* (Oxford University Press, New York).

Small ML, Jacob EM, Massengill RP (2008) Why organizational ties matter for neighborhood effects: Resource access through childcare centers. *Soc. Forces* 87(1):387–414.

Snijders TAB, van de Bunt GG, Steglich CEG (2010) Introduction to stochastic actor-based models for network dynamics. *Soc. Networks* 32(1):44–60.

Valente TW, Coronges K, Lakon C, Costenbader E (2008) How correlated are network centrality measures? *Connections (Toronto)* 28(1):16–26.

Wimmer A, Lewis K (2010) Beyond and below racial homophily: ERG models of a friendship network documented on Facebook. *Amer. J. Sociol.* 116(2):583–642.