

## Factors of Students' Network Formation: Case of a Russian University

Krekhovets E.,

*National Research University Higher School of Economics, Researcher, Lab for Institutional Analysis,  
Centre for Institutional Studies, Higher School of Economics, Moscow, Russia.*

Larin A.,

*National Research University Higher School of Economics*

Doi:10.5901/jesr.2013.v3n7p647

### Abstract

*This paper investigates the process of students' network formation in National Research University–Higher School of Economics. Using pairwise regression, estimated on a network of second-year students from two faculties, we find that geographical closeness along with the homophily effects of gender, grade point average, smoking, and leisure preferences have a strong impact on the probability of being friends.*

**Keywords:** social network, higher education, homophily effect, geographical closeness

### 1. Introduction

In recent decades, social network analysis (SNA) has grown into a popular method applied to various social researches. SNA is used in various socioeconomic researches, including economics of education. SNA is believed to be one of the relevant lines of investigation of students' behavior and academic achievements in the university. Instruments of SNA allow for exploring new factors that influence behavior and academic performance and give an opportunity for the complete description of the social capital formation process. SNA is consistent with a notion of university, which is known to be a place where people form their relations. Furthermore, SNA allows the broadening and enriching of social understanding in a higher education institute (Biancani, 2013).

The majority of modern SNA studies on students' social networks tend to focus on two main aspects: (1) which factors determine network formation and (2) how these networks influence people's behavior. When investigating the role of networks, researchers estimate different types of peer effects that show how network characteristics may influence individual outcomes. When examining which factors lead to student network formation, researchers document factors such as race and ethnicity, socioeconomic characteristics, age, cultural habits, and cohorts. In the literature, there are consistent evidences supporting that students prefer to be friends with students similar to themselves. This effect is known as homophily. Physical and social propinquity is essential as well. If students are roommates in the dorm or they live geographically close to one another or they are enrolled in the same educational program, the probability that they will become friend increases.

In this paper, we try to examine factors which determine students' social network formation in National Research University–Higher School of Economics (HSE) in Nizhny Novgorod. For this study, 237 students from two faculties—the faculty of economics and the faculty of management—were interviewed. The questionnaire includes questions about students' academic performance and behavior, their attitude toward friendship, their socioeconomic characteristics, their hobbies, and their leisure time.

Using information from about 24,090 possible links, we estimate pairwise regression and present the marginal effects on the probability of being friends. We find that geographic closeness is one of the most influential factors of the network formation process. Data also support the homophily hypotheses for gender, grade point average, smoking, and leisure preferences, but not for income and alcohol. We do not find any signs of the impact of study preferences.

The rest of the article is organized as follows. In Section 2, several literatures focusing on students' network formation are briefly reviewed. The data and structure of students' network are examined in Section 3. Findings and results are presented in Section 4. We conclude in Section 5.

## 2. Determinants of Students' Network Links: Literature Review

A network can be presented as a set of collective and individual links which unite people into social groups (Pattison, 1993). The concept of a social network is used for describing social relationships on different levels, from interpersonal to international. Formally, any network is a set of nodes and links between these nodes. In social networks, nodes are usually persons, groups of people, companies, or counties. Friendship, kinship, common interests, and business relations constitute links in the social network. Therefore, a social network defines social interrelations, and the study of social networks expands the set of factors that can explain human behavior (Newman, 2010).

In one of the first investigations of factors that determine students' social networks, Festinger, Schachter, & Back (1950) studied student couples from a U.S. university. Authors found a strong propinquity effect: couples that lived geographically close were more likely to become friends.

The majority of literature on students' social network formation examines the role that race and ethnicity play in the formation of friendship among students. Kenny & Stryker (1994) compared several groups of students of different races. They concluded that students' social networks were created under the influence of race and ethnicity. This effect is also known as racial homophily. Moreover, researchers compared characteristics of networks formed by students of different races. As a result, they found that white students organize more extensive and dense networks than their nonwhite classmates. Race and ethnicity how students adapt to college: white students receive support from their friends, while students of other races prefer to communicate more with their parents and relatives.

Mayer & Puller (2008) used Facebook.com to gathered data about students from 10 universities in Texas. They developed a model of student social network formation that found two main groups of influential factors. The first one is a set of exogenous factors which reflect features of an academic environment. The second group of factors consists of endogenous factors linked with individual student's characteristics. According to the research, race is considered to be an endogenous factor, and the policies set by a university have very little influence on the separation of social network by race. On the whole, they found that students' social network formation is determined by race, cohorts, majors, and university policies.

Wimmer & Lewis (2010) also used data from Facebook.com and found the noticeable influence of racial homophily on social network formation among students. The study supported the idea that racial homophily is an influential but not dominant mechanism that determines students' friendship. It is noteworthy that socioeconomic characteristics play an important role in this. Among other factors, authors emphasize choice of the same academic major, reciprocation, musical and cultural homophily, and propinquity. Some of these factors are likely to be more essential than race and ethnicity.

Contrary to Mayer & Puller (2008), Wejnert (2010) came to the conclusion that university policies have an influence on the friendship of students of different races. One way to do so is through the accommodation of students to racially diverse dormitories. Living with roommates of different races increases the likelihood of creating interracial friendships. Notably, these ties prompted by university policies are as strong as natural ones formed randomly.

According to Marmaros & Sacerdote (2006), among the key factors that determine the formation of students' networks are ethnicity and race, family background, geographical closeness, and common interests in university, for example, choice of a similar major as well as participation in sports teams or Greek organizations. Using the data from Dartmouth College, authors found that racial similarity and geographic closeness are more influential than family background, common academic major, and playing in varsity. The mechanism of racial homophily had proved to be very influential: two white students have twice as many chances to become friends with each other than with students of a different race. It should be noted that living in the same dormitory has no positive effect on interracial friendships. In contrast, different-race students educated in the same class are likely to become friends.

Using data from Facebook.com, Traud, Kelsic, Mucha, & Porter (2011) analyzed which factors play significant roles in social network formation in five geographically diverse universities in the United States.

Findings from several other studies support the idea that geographical closeness is a crucial factor that affects students' social networks. Lee, Scherngell, & Barber (2011) used the data collected from the online social network platform StudiVZ about several German universities to estimate the spatial interaction model. They found that geographical closeness has a direct effect on students' social network formation. According to their estimates, the distance equal to 100 minutes by car reduces the probability of social network formation by 91%. They also found that students are more likely to become friends with those who studied at the university located at the same federal state. Additionally, results showed that students' social network formation depends on university programs: students from technical and mathematical universities prefer creating ties with those who are enrolled in the same courses. The

findings showed that living in the same dormitory, studying in the same year and class, and demographic characteristics are likely to determine friendship among students. Researchers came to the conclusion that there is no dominant factor in friendships and that students' social networks form under the combined influences of several factors.

Therefore, the majority of the studies focused on two main mechanisms of students' social network formation: propinquity and homophily. In the literature, there is conclusive proof that students are likely to become friends with those who are of the same race and ethnicity, have equal socioeconomic status, and have the same interests. Additionally, the choice of the same academic major and participation in varsity and students' organizations have a positive impact on the likelihood of the formation of ties among students. Roommates also have a better chance of becoming friends, and living with different-race students has a significant influence on the formation of interracial students' social network. It means that university policies are able to effect the formation of different-race social networks. In addition to the above-mentioned factors, geographical distance is essential: the closer students live to each other, the more likely they will become friends.

### 3. Data Description

Data for this analysis are gathered from the questionnaire survey of second-year students of HSE in Nizhny Novgorod, Russia. The sample used in this study includes 220 students from 10 classes (5 classes from the faculty of economics and 5 classes from the faculty of management). The total number of students enrolled in the faculties of economics and management is 237. Students were asked 69 questions about their social status, academic environment, relationships within classes, and leisure time. We also used administrative information about the students' grade point averages (GPAs). There is an open 10-point assessment system in the university, and students are able to know the grades of their peers. Descriptive statistics are presented in Table 1.

**Table 1.** Descriptive Statistics

	Mean	Min.	Max.	No. of observations
<b>Faculty of Economics</b>				
Age	19 (0.67)	17	20	99
GPA	7.09 (1.08)	4.56	9.33	99
% of men	26			99
% of tuition free students	80			99
<b>Faculty of Management</b>				
Age	19 (0.62)	17	20	121
GPA	7.56 (0.81)	5.28	9	121
% of men	31			121
% of tuition free students	69			121

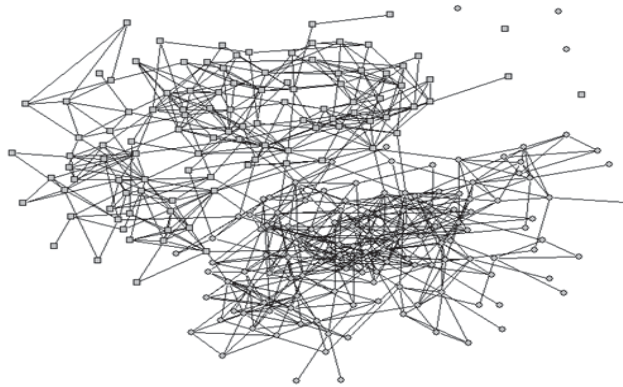
The descriptive statistics of main students' characteristics, which are used in the model, are presented in Table A.1 in the appendix. According to statistics, about half of the students in the sample have a personal car (57%). More than 27% of the students live in a dormitory. Other students live with their parents, in their personal apartments, or in apartments for rent. About one-third of the students have a per-capita family income between 6,400 and 15,000 rubles. Roughly 44% of students prefer humanitarian courses, about 25% of students prefer mathematical and applied courses (27% and 24%, respectively), and only 5% of students mention theoretical courses as their preferences. Students mostly communicate only in the university (67% of respondents) and participate in the social life of their university. About 62% of students attend cultural events like theaters, cinemas, and exhibitions. But only 16% go to nightclubs. The majority of respondents have no pernicious habits: 12% are smokers, 5% are frequent drinkers of alcohol, and 37% are athletic.

Using the Pajek package, we create a directed graph of the network, which is presented in Figure 1. The main network characteristics are presented in Table 2.

**Table 2.** Network Characteristics

Characteristics	Value
Density	0.0186
Average degree	8.8101
Diameter	15.000
Average distance among reachable pairs	5.2976
Network clustering coefficient	0.2512
Watts-Strogatz clustering coefficient	0.3023
Network betweenness centralization	0.0975

The network is not connected. The size of the largest component of the network is 232 vertices, more than 97% of the total network. There are 5 students (2 from the faculty of economics and 3 from the faculty of management) who did not note other students as friends, and other students also did not note them. According to statistics, the diameter is 15. This characteristic shows the longest path between nodes in the networks. A short diameter means that students in two faculties are sufficiently close. The average degree value shows that students generally have about 9 friends in their university environment. One more network property is the clustering coefficient, which argues that about a quarter of the friends are friends of one another. Betweenness centrality refers to the importance of students as connectors.



**Figure 1.** Students’ network graph

Figure 1 gives a clear notion of student clustering—there are two distinct regions in the graph (with a small number of links between them) that represent the two faculties. Students’ groups are not so clearly defined since they have a lot of links within the faculties.

#### 4. Results

In this section, we present estimates of how factors such as geographic closeness, sociodemographic characteristics, and preferences affect the formation of friendship among students. To obtain these estimates, we use pairwise regression (Fafchamps & Gubert, 2007; Mayer & Puller, 2008; Udry & Conley, 2005). The idea is to consider not the students themselves, but the links between them. This approach is based on the binary choice model:

$$y_{ij} = \begin{cases} 1 & \text{if } y_{ij}^* \geq 0, \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

where  $y_{ij}$  denotes the existence of a link between nodes  $i$  and  $j$ . Latent variable  $y_{ij}^*$  denotes the propensity to form that link and depends on the vector of observable variables  $x_{ij}$  and unobservable shock  $\epsilon_{ij}$  in the following manner:

$$y_{ij}^* = x_{ij}'\beta + \epsilon_{ij}. \tag{2}$$

Explanatory variables  $x_{ij}$  may contain student-specific characteristics (GPA, age, some preferences, and so on) and some characteristics that refer to both of the students  $i$  and  $j$  (neighborhood, common friends, and so on). Here  $\beta$  is the vector of parameters to be estimated and reflects the impact of  $x_{ij}$ .

We assume that unobservable shock  $\varepsilon_{ij}$  follows logistic distribution and use the maximum-likelihood estimator to obtain estimates of unknown parameters. The interpretation of parameters  $\beta$  is complicated—they show how  $x_{ij}$  affects latent variable  $y_{ij}^*$ , but not  $y_{ij}$ . Thus, estimates of  $\beta$  only reveal the sign of the impact of  $x_{ij}$  on  $y_{ij}$  and whether this impact is significant. That is why we present average marginal effects that show how a change in  $x_{ij}$  affects the probability of students  $i$  and  $j$  to become friends.

We use undirected notation of the network and mark that the link between students  $i$  and  $j$  exists if at least one of them says that they are friends. Since we have 220 students in the network (with available data), there are 24,090 possible links between them, and thus, we have 24,090 observations in the sample. But only 669 of that links are realized. In other words, only 669 of 24,090 possible pairs of students are marked as friends.

Estimation results are presented in Table 3. There are 7 panels in the table that refer to different groups of explanatory variables.

To investigate the impact of gender (*Panel A*), we use two dummy variables. *Both female* equals 1 only if the observation is of a pair of female students and 0 otherwise. *Both male* equals 1 only if the observation is of a pair of male students and 0 otherwise. Thus, a pair of students of different sexes is a base category. The estimates say that the probability of friendship between students of the same sex is higher—by 0.4% for female students and by 1.2% for male students—in comparison with students of different sexes. One may suggest that girls have fewer incentives to be friends with students of the same sex since estimates for them are lower. But this result is probably explained by the fact that the share of female students is much larger (74%) and thus, the probability of friendship between two random girls is lower than between two random boys.

**Table 3.** Marginal Effects on the Probability of Being Friends

Variable	Marginal Effect (%)	Variable	Marginal Effect (%)
<i>Panel A. Gender</i>			
Both female	0.356*** (0.063)	Both male	1.162*** (0.264)
<i>Panel B. Neighborhood</i>			
Live in dormitory, both	0.360*** (0.135)	Live in dormitory, none	0.149*** (0.047)
Roommates in dormitory	31.364*** (9.701)		
Same group	6.455*** (0.879)		
Same faculty	1.174*** (0.120)		
<i>Panel C. Socioeconomic Characteristics</i>			
Income, difference	-0.022 (0.016)	Income, average	0.060*** (0.023)
Have car, both	0.174 (0.127)	Have car, none	0.077* (0.041)
<i>Panel D. Grade</i>			
GPA, difference	-0.136*** (0.032)	GPA, average	0.112*** (0.036)
Budget place, both	-0.001 (0.045)	Budget place, none	-0.032 (0.077)
<i>Panel E. Study Preferences</i>			

Like mathematics, both	0.026 (0.076)	Like mathematics, none	-0.040 (0.048)
Like humanities, both	0.014 (0.055)	Like humanities, none	0.028 (0.053)
Like theoretical, both	-0.061 (0.298)	Like theoretical, none	0.099 (0.069)
Like applied, both	-0.043 (0.075)	Like applied, none	-0.016 (0.044)
<b>Panel F. Social Life and Communication</b>			
Rest with friends, both	0.173** (0.077)	Rest with friends, none	-0.038 (0.041)
Like sports, both	0.116* (0.069)	Like sports, none	0.041 (0.042)
University social life, both	0.041 (0.041)	University social life, none	0.021 (0.063)
Cultural events, both	-0.052 (0.039)	Cultural events, none	-0.003 (0.058)
Nightclubs, both	0.321* (0.166)	Nightclubs, none	0.085* (0.044)
Often ill, both	0.011 (0.222)	Often ill, none	0.029 (0.05)
<b>Panel G. Habits</b>			
Alcohol, difference	-0.025 (0.027)	Alcohol, average	-0.031 (0.038)
Smoke, both	1.027** (0.478)	Smoke, none	0.091* (0.049)
<b>Number of observations</b>	<b>24 090</b>		

**Notes:** Standard errors in parentheses. \*\*\*, \*\*, \* denote significance at 1%, 5%, and 10% levels, respectively.

*Panel B* refers to the impact of geographic closeness. It says that if students are roommates in a dormitory, then the probability of friendship between them is higher by 31.4% in comparison with students who are not neighbors. The fact that students study in the same faculty/group also raises this probability. One may note that the effect of geographic closeness depends on the size of the area in question. For example, the size of a group is five times smaller than the size of a faculty, and the effect of the *same group* is five times larger than the effect of the *same faculty*.

The fact that both of the students live or do not live in dormitories also raises the probability of friendship. But this result can be explained not only by geographic closeness but also by the effect of homophily. Students who live in dormitories are more likely to make friends with those who have the same socioeconomic characteristics—nonresidents who do not rent a flat.

*Panel C* reflects the direct effect of socioeconomic characteristics. We use two variables that are constructed on the basis of family per capita income of the students. In order to catch the homophily effect, we include the absolute value of *income difference*, which reflects how the income of student  $i$  differs from the income of student  $j$ . And we include *income average* to reveal how income may be connected with students' sociality. Data do not support the hypothesis of income homophily as the marginal effect of *income difference* is not significant. But we may say that students with higher family incomes tend to have a higher probability of being friends.

At the same time, *Panel D* shows that a 1-point difference in GPA lowers the probability of friendship by 0.1%. And analogously to income, a higher GPA is connected with a greater probability of being friends.

*Panel E* does not report any significant impact of *study preferences*. A possible reason for this result is that preferences are not varied enough since all the students are from the same university. In this case, there is no contrast between mathematics and humanities, for example, and it is hard to reveal precise estimates of the effects of such preferences. Moreover, this effect may be partly captured by faculty and group dummies.

*Panel F* shows that the probability of friendship is raised if both of the students *communicate with classmates* outside the university or participate in *sports* or frequent *nightclubs*. The effect of communication with classmates is probably connected with students' sociality, while the last two effects may also be caused by preference homophily. This hypothesis is supported by the fact that the friendship probability is raised even if both students do not visit nightclubs. In other words, students who do not visit nightclubs have little incentive to make friends with those who do so.

Panel G does not show any impact of alcohol but reports the significant effect of smoking—smokers are more likely to be friends with smokers and nonsmokers with nonsmokers.

## 5. Concluding Remarks

In the paper, we investigate the process of students' network formation in HSE in Nizhny Novgorod. Using data on second-year students from two faculties, we estimate pairwise regression and obtain the marginal effects of factors such as geographic closeness, sociodemographic characteristics, and students' preferences.

We find that geographic closeness has the most considerable effect: if students become roommates in a dormitory, then the probability of friendship between them rises by 31%. The impacts of the same faculty and the same group are also significant.

Though we do not find any signs of impact of income gap, results are in favor of homophily through factors such as gender, grade point average, smoking, and leisure preferences (sports and nightclubs). We also show that the probability of friendship is positively connected with the average family income and grade point average of students. This may be caused by the popularity of students and/or their sociality.

In contrast to recent researches, we do not find a significant impact of study preferences. We suppose that these results are not robust and are caused only by the fact that all students are from the same university and related faculties, that leads to the low heterogeneity of study preferences. In this connection we expect that a significant impact may be found when other universities/faculties are added to the sample.

## References

- Biancani, S., (2013). Social Networks Research in Higher Education. *Higher Education: Handbook of Theory and Research*, 28, 151-215.
- Fafchamps, M., & Gubert, F. (2007). Risk Sharing Networks in Rural Philippines. *Journal of Development Economics*, 71, 261–287.
- Festinger, L., Schachter, S., & Back, K. (1950). *Social Pressure in Informal Groups*. Stanford, CA: Stanford University Press.
- Kenny, M., & Stryker, S. (1994). Social Network Characteristics of White, African-American, Asian and Latino/a College Students and College Adjustment: A Longitudinal Study. In *102nd Annual Meeting of the American Psychological Association, Los Angeles, CA, USA*.
- Lee, C., Scherngell, N., & Barber, M. J. (2011). Investigating an Online Social Network Using Spatial Interaction Models. *Social Networks*, 33(2), 129–133.
- Marmaros, D., & Sacerdote, B. (2006). How Do Friendships Form? *The Quarterly Journal of Economics*, 80, 79–119.
- Mayer, A., & Puller, S. (2008). The Old Boy and Girl Network: Social Network Formation on University Campuses. *Journal of Public Economics*, 92, 329–47.
- Newman, M. E. J. (2010). *Networks: An introduction*. Oxford, UK: Oxford University Press.
- Pattison, P. (1993). *Algebraic Models for Social Networks*. New York, NY: Cambridge University Press.
- Traud, A. L., Kelsic, E. D., Mucha, P. J., & Porter, M. A. (2011). Comparing Community Structure to Characteristics in Online Collegiate Social Networks. *SIAM Review*, 53(3), 526–543.
- Udry, C., & Conley, T. (2005). Social Networks in Ghana. In C. B. Barrett (Ed.), *The Social Economics of Poverty: Identities, Groups, Communities and Networks*. London: Routledge, chap. 10.
- Wejnert, C. (2010). Social Network Analysis With Respondent-driven Sampling Data: A Study of Racial Integration on Campus. *Social Networks*, 32(2), 112–124.
- Wimmer, A., & Lewis, K. (2010). Beyond and Below Racial Homophily: ERG Models of Friendship Network Documented on Facebook. *The American Journal of Sociology*, 116(2), 583–642.

**Table A.1.** Detailed Descriptive Statistics of Students' Characteristics

Per Capita Family Income							
	Less than 6,400	6,400–15,000	15,000–21,000	21,000–30,000	30,000–50,000	50,000 and more	No. of observations
% of Students	15	30	20	12	15	8	220
Place of Living							
	Living in dormitory		Living in other place				No. of observations
% of Students	27.3		72.7				220
Personal Car							
	Yes		No				No. of observations
% of Students	57		43				220
Study Preferences							
	Mathematical	Humanitarian	Theoretical	Applied		No. of observations	
% of Students	27	44	5	24		220	
Communication with Classmates							
	Out of the university		Only in the university				No. of observations
% of Students	33		67				
Sports Activity							
	Yes		No				No. of observations
% of Students	37		63				220
Participation in University Social Life							
	Yes		No				No. of observations
% of Students	63		37				220
Attendance in Cultural Events							
	Yes		No				No. of observations
% of Students	62		38				220
Attendance in Nightclubs							
	Yes		No				No. of observations
% of Students	16		84				220
Smoking							
	Yes		No				No. of observations
% of Students	12		88				220
Drinking Alcohol							
	Frequently	Sometimes	Rarely	Never		No. of observations	
% of Students	5	17	39	39		220	