How Social Ties Affect Peer Group Effects: Case of University Students[1]

Oleg Poldin, Dilyara Valeeva, Maria Yudkevich

National Research University - Higher School of Economics

Among the key issues in the empirical estimation of peer effects are correct identification of relevant peers and endogeneity of their outcomes ("reflection problem"). In this study, individual university student performance is related to the performance and the characteristics of the person's social network members. The analysis uses data on two directed networks: friends' network and study partners' network of third-year students at one Russian top-tier research university. Data on network ties enables us to address "reflection problem" and disentangle influence of peers' outcome from effect of peers' background characteristics. The findings are that both peers' GPA and peers' ability measures are significant in the estimated regression model. One point increase in peer mean GPA is associated with an increase in own GPA of approximately one fourth. The regression with study partners' network data has slightly greater explanatory power than the analysis based on friendship network data. No effects from student classmates are found in a model that assumes group interactions in student micro-groups.

1. Introduction

The result of the individual learning is affected by the wide range of factors. Some of factors are in the sphere of traditional research interest because they are inputs of the educational production function, e.g., student's characteristics (including the personal ability), teacher quality, school resources, and peer characteristics. The peer characteristics and behavior can also play an important role in educational achievement. This influence is called peer effect. While there is much research devoted to the estimation of peer effects in secondary schools, studies of these effects in higher education are not so abundant.

One strand of empirical work on peer effects in higher education institutions is based on the analysis of data about students living in one room or one section in student dorms. Sacerdote (2001), Zimmerman (2003), Brunello et al. (2010) use randomly assigned roommates as peers. Sacerdote (2001) found that average grades were higher for those students whose roommate was in the top 25%

[1] Authors thank all participants of Center for Institutional Studies research seminar for valuable comments and fruitful discussion. Financial support of Basic Research program (HSE) is greatly appreciated.

of the class. In Zimmerman (2003) study, students in the middle of the SAT score distribution get worse grades if their roommates are students with low grades. . In Brunello et al (2010), positive and significant effects are found for students specializing in engineering and mathematics.

The other approach relies on an entire study group as a peer group. However, free choice of subjects to study in most universities implies endogenous formation of study groups, and this impedes correct estimation of peer effects. One of exceptions to this practice happens in military institutions. In the specific environment of these institutions, students interact intensively within the groups formed administratively. Lyle (2007) finds a significant relationship between achievement of first year students and average achievement of the group in the US Military Academy. It is also revealed that the increased dispersion of math SAT scores in a group improved student achievement, and that the given effect was achieved due to the presence of more able students (Lyle 2009). Carrell, Fullerton, and West (2009) find significant peer effects among students of the US Air Force Academy: students with low verbal SAT grades benefit from their communication with students with high SAT results. De Paola and Scoppa (2010) report on significant peer effects for the University of Calabria in Italy. Arcidiacono, Foster, Goodpaster, and Kinsler (2012) find statistically significant peer effects on course grades, particularly in courses of a collaborative nature.

In some countries, higher education institutions practice exogenous formation of student groups and have curricula dominated by compulsory courses. Androuschak et al. (2012) estimate the influence of classmates' ability on first-year student's achievement at a top-tier research university in Russia. The presence of high-ability classmates has a significant positive effect on individual grades in key economic and mathematical courses as well as overall academic performance. Students at the top of the ability distribution derive the greatest benefit from high-ability classmates. Less able students are not affected by peers and have no significant influence on peers' outcomes.

Study of the peer effects from the entire group has certain drawbacks. Thus, it's supposed that all the members of a group have the same influence on a student. If this assumption is quite possible in the first year of study, it fails later because students form their own specific social ties. Almost all the previous work on peer effects in higher education focuses on the first year, so the study of how social interactions affect student achievement beyond the first year is of obvious interest.

Also, the group interaction assumption complicates differentiation between the effects of peer characteristics and the effect of peer outcomes. This differentiation is very important both for the understanding the mechanisms of peer group effects and for the educational policy. Some characteristics of the peers (abilities, socio-demographic characteristics) are fixed and cannot be changed. Peer behavior (achievement) can be changed by the influence of educational policy. For example, if peer outcomes are matter, then stimulation of achievement of a target part of students affect achievement of their peers, then the peers affect their peers (including target students), and so on, thus creating the effect of social multiplier.

Analysis of each student's individual connections allows researchers to overcome above-mentioned limitations imposed by assumption of group interaction. Recently, several studies addressed the effect of social ties on the behavior and achievement of students.

Lavy and Sand (2012) study the influence of different types of friendship relationships on achievement and behavior of schoolchildren in Israel. Their results suggest that the presence of reciprocal friends (students who list one another) and followers (those who listed fellow students as friends but were not listed as friends by these same fellow students) in class has a positive and significant effect. Empirical research by Calvó-Armengol et al. (2009), Bramoullé et al. (2009) and Lin (2010) is based on the extensive Add Health database on high school students in the USA that contains information about friendship relations between respondents. Calvó-Armengol et al. (2009) show that the number of friends is a positive significant factor of academic achievement. Bramoullé et al. (2009) show that the mean level of recreational activities (participation in artistic, sports and social organizations and clubs) of friends have a positive and significant influence on a student's recreational activity. Lin (2010) finds that both achievement of friends and their social-economical characteristics significantly correlated with student's own achievement.

The above papers on peer effects in social networks use data on students in the secondary schools. Studies of the network impact on student achievements in university environment are mostly limited to data from Internet activities (chats, on-line social networks) or from an e-mail exchange. Mayer and Puller (2008) use data generated from student accounts in the Facebook social network of one US university. Among other results, they find that performance of students is highly correlated with the average performance of their friends: increase of one point in the friends' GPA increased own average grade of 0.46. However, web data should be used with caution since they give rather raw estimates of real world friendship ties.

In this work we use data on social connections of third year students of one of the Russian universities. Data was gathered by the questionnaire survey. Besides friendship ties, data on ties with study helpers are used. We find significant positive peer group effects both from academic achievement and from peer abilities. The influence of groupmates as a whole is insignificant. The results of this study help to better understand actual mechanisms of peer effects in group interactions.

2. Empirical models of peer effects

We start from the following specification of the empirical peer effects model:

$$y_i = \alpha + \beta \overline{y}_{-i}^{peer} + \boldsymbol{\gamma}'\mathbf{x}_i + \boldsymbol{\delta}'\overline{\mathbf{x}}_{-i}^{peer} + \varepsilon_i \,,\; i = 1,2,...,n., \tag{1}$$

where y_i is an indicator of academic achievement of student *i*, $\mathbf{x}_i$ is k ×1 vector of individual characteristics of student *i*, $\overline{\mathbf{x}}_{-i}^{peer}$ is k ×1 vector of mean exogenous characteristics of peer group for student *i*, $\overline{y}_{-i}^{peer}$ is the mean academic achievement in peer group of student *i*, ε_i are random disturbances, *n* is the number of students in the sample.

Coefficient β in model (1) measures the degree to which an individual achievement depends on achievement of peers (endogenous effect). Components of a vector $\boldsymbol{\delta}$ measure the influence of exogenous characteristics of peers (exogenous, or contextual effect). The important difference between these impacts is that an endogenous effect only ($\beta \neq 0$) has social multiplier property. Indeed, suppose that achievement of a certain student becomes higher as a result of some external influence (ε_i in the equation (1)). This improvement makes an influence on his or her peers' performance. In turn, a change in achievement of the peers also has a feedback influence on a student, etc. Pure exogenous peer effects ($\boldsymbol{\delta} \neq 0$, $\beta = 0$) don't have a multiplier effect because external shocks don't influence fixed characteristics of students.

Presence of endogenous effects is significant for educational policy. For example, additional training with certain students can lead to better achievement of not only those students that attend these additional classes, but can also cause the increase in the attainments of their peers.

Since not only achievement of the student's peers have an effect on his or her academic outcomes but also his or her own performance influences peers outcomes, estimates of the coefficients of the model (1) obtained by the least square method are biased. To avoid this bias, the reduced form model is frequently used:

$$y_i = \alpha + \boldsymbol{\gamma}'\mathbf{x}_i + \boldsymbol{\delta}'\overline{\mathbf{x}}_{-i}^{peer} + \varepsilon_i \quad . \tag{2}$$

Estimates of reduced form can answer the question on presence of peer effects, but of no help with the differentiation of endogenous and exogenous effects.

However, one may estimate separately endogenous and exogenous peer effects if data about student's personal social ties are available.

Assume that it is students connected one to another with social ties who influence academic achievement. Each student has his or her own social network. Then we can write the peers mean values from (1) this way:

$$\overline{\mathbf{x}}_{-i}^{peer} = \frac{1}{n_i}\sum_{j=1}^{n_i}\mathbf{x}_j, \quad \overline{y}_{-i}^{peer} = \frac{1}{n_i}\sum_{j=1}^{n_i}y_j, \tag{3}$$

where averaging is over set P_i which consists of n_i peers linked to a student *i*. When the means are calculated, own characteristics of a student are not included.

Then the model (1) for the sample of n students can be written in matrix form:

$$\mathbf{y} = \alpha\mathbf{i} + \beta\mathbf{Gy} + \mathbf{X}\boldsymbol{\gamma} + \mathbf{GX}\boldsymbol{\delta} + \boldsymbol{\varepsilon},\quad E\left[\boldsymbol{\varepsilon} \mid \mathbf{X}\right] = 0, \tag{4}$$

where **y** is n ×1 vector of achievement measure of n students , **i** is n ×1 unit vector, **X** is n ×k matrix of characteristics of students, **G** is n ×n matrix of interactions between elements

$$G_{ij} = \begin{cases} 1/n_i, & если\ j \in P_i, \\ 0, & если\ j \notin P_i. \end{cases}$$

G is the row normalized matrix with zero diagonal elements. Row i of matrix **GX** consists of mean values of k characteristics of peers of a student i, and element i of vector **Gy** is mean achievement of peers of a student i:

$$\mathbf{GX} = \begin{pmatrix} \bar{X}_{-1}^{peer} \\ \vdots \\ \bar{X}_{-n}^{peer} \end{pmatrix},\quad \mathbf{Gy} = \begin{pmatrix} \bar{Y}_{-1}^{peer} \\ \vdots \\ \bar{Y}_{-n}^{peer} \end{pmatrix}. \tag{5}$$

Model (4) structurally looks like an autoregression model of spatial interaction (LeSage, Pace (2009) classifies this model as Spatial Durbin Model) that can be estimated by the maximum likelihood method. In spatial statistics matrix **G** is similar to the matrix of spatial weights that describes the geographical proximity between different objects in a space and therefore is symmetrical. In case of the social ties matrix **G** can be non-symmetrical if student i nominates student j as a friend but not vice versa.

The problem of estimating vector parameters α, β, $\boldsymbol{\gamma}$, $\boldsymbol{\delta}$ in (4) is related to the endogeneity of **GY** on the right side. As Bramoullé, Djebbari and Fortin (2009) showed, the parameters can be identified if matrices **I**, **G**, and $\mathbf{G}^2$ are linearly independent. In network interactions, linear independence is assured by the presence of intransitive triads in a network structure. The triad of students A, B and C is intransitive if A affects B and B affects C but A does not affect C. This requirement is usually holds in case of social ties because usually not all friends of the friends are also the friends of a student.

As mentioned above, change of the student characteristics affects both student own achievement (direct impact) and an achievement of other students (indirect impact). Estimations of the coefficients of spatial (network) regression take into account an interdependence between observations for a particular student. These characteristics advantageously differentiate this model from a simple regression model where observations are independent from each other.

Some quantitative measures that characterize connection between dependent and independent variables and reflect the specificity of network interactions are described in Appendix. Average direct impact characterizes an impact of student's characteristics on his or her own outcome averaged over the sample. Average total impact measures an influence on achievement of one student from characteristics of all other students or, equivalently, measures the influence from a student's characteristics on all other students. Average indirect impact, by the definition, is a difference between average full impact and average direct impact. In section 3 we present estimations of these parameters for the observed network.

3. Estimation of peer effects in a social network

Data

In this study, we use data on academic achievement and characteristics of students Economics department students in the Higher School of Economics, which studied at the 3rd year in 2011-2012 academic year. As a measure of academic achievement we use student grades averaged over three years. In the Higher School of Economics, a 10 point grade system is adopted. The higher the grade, the higher is the achievement on a certain subject.

The student cohort consists of 231 students divided into 8 student groups formed prior to beginning of first academic year for first three years of study by the university administration. Lectures are usually delivered to several groups simultaneously, while seminars and tutorials (classes) are delivered to each group separately. Therefore, the HSE students spend one part of their study time all together, another part of time the students study in groups.

For the correct determination of each student's peer group, we used a questionnaire survey in the mid of the third year of study. Students were asked to indicate no more than five students with whom they usually spend their free time ("friends") and no more than five students to whom they address for some educational help ("study partners"). The data sample used is smaller than the number of all enrolled students. Some students missed the classes when the survey data were collected. Also, we exclude the students who were not nominated as friends and study partners.

After processing the data from the questionnaire, we formed two types of matrices of social relations of each student, one matrix for friends and another one for study assistants. Proportion of friends among study assistants is 60 %.

Most important predictor of student's academic achievement in college is personal ability measured by previous achievement (Burton and Ramist, 2001, Noble and Sawyer, 2002, Bauer and Liang, 2003, Shaw et al., 2012). In our study, as an exogenous measure of individual student ability we used the grades on Unified State Examination (USE) in Mathematics and Russian language. The USE in Mathematics and Russian language are national standardized tests, they are compulsory for all secondary school graduates and used as sole basis for university admission decisions. Besides the test scores, we measure ability by dichotomous variables which are indicators of awardees of All-Russian

Olympiads and regional Olympiad for secondary school students. Awardees of Olympiads have right of priority in admission to higher education institutions in Russia.

Although the the HSE students come from different regions of the country, ethnic composition is homogeneous so there is no need to introduce ethnic dummies. Descriptive statistics are presented in Table 1.

In Table 2 some network characteristics are presented. All the values are measured from 0 to 1. Help network is more centralized than friendship network. We can explain this by the fact that students usually tend to address to the same high-achieving students. Friendship network is more reciprocal than help networks; almost half of all the possible ties are mutual.

Estimation results

We estimate both the model in a reduced form without endogenous term and the model in a full specification that differentiates between endogenous and contextual peer effects. In addition to full set of the exogenous parameters characterizing abilities (USE scores on Math and Russian language, status of Olympiad winners and awardees), we also consider specifications with only USE scores as exogenous variables.

Table 3 reports the least squares estimates of the reduced form model for friendship ties:

$$\mathbf{y} = \alpha_r \mathbf{i} + \mathbf{X}\boldsymbol{\gamma}_r + \mathbf{G}\mathbf{X}\boldsymbol{\delta}_r + \boldsymbol{\varepsilon}_r. \tag{6}$$

Own characteristics of a student are significant at 1% level. The influence of USE in Russian language is weaker than the influence of USE in Math: the increase in USE in Math on 10 points leads to the increase of GPA of 0.41, the increase in USE in Russian language leads to the GPA increase of 0.23. The status of awardee of All-Russian Olympiad considerably increases GPA (0.81 points) that is larger than status of awardee of the Regional Olympiad of the HSE (0.43 points).

In this model, coefficient $\boldsymbol{\delta}_r$ “mixes” endogenous and exogenous effects, not allowing identifying them separately. The mean USE Math scores of friends and the mean number of friends that are awardees of All-Russian Olympiad are significant variables. Thus, there is an evidence of peer effects: own achievement of a student increases when USE Math scores of his or her friends grow as well as when this student has friends who are awardees of All-Russian Olympiad.

The natural assumption is that strong friendship ties involve reciprocity. In columns 3 and 4 of table 3, we report the estimates for a case with reciprocal friendship ties: two students are considered to be friends if their nominations are mutual. For reciprocal ties, the friends’ USE in Math is significant at the 1% level.

Estimates of the model in a full specification spatial autoregression model (4) using the method of maximum likelihood are shown in Table 4a. The values of the coefficients β are significantly positive. The increase in the GPA of friends of one point is associated with the increase of own student's GPA of 0.25. The coefficients of exogenous variables are not significant for directed ties. In case of reciprocal ties, the mean value of friends' USE in Math is significant at the 10% level. This result supports assumption that the success of a student is influenced by actual achievement of the friends rather than by their exogenous characteristics.

In Table 4b, estimates of average direct, average indirect and average total impact of exogenous characteristics of the friends are presented.

Students that are connected by friendship ties usually study at the same student group. To compare peer effects generated by friends and by the entire study group, we estimated models (1) and (2) for the situation when the peer group supposed to include all students of a group. The results are presented in Table 5. Influence of exogenous characteristics of a student group is positive but weaker than the influence caused by friends. Both exogenous and endogenous effects are statistically insignificant.

In the questionnaire besides nomination of the friends, students nominated students to whom they address for help in studying. The lists of friends and study partners are partly overlapping but generally for an obvious reason students tend to consult on educational questions with the fellows that have higher achievement.

In Table 6a, we present the results for study partners' networks. The results are nearly similar to the results from Tables 3 and 4 for friendship network. The coefficient of determination for help network model is a little bit higher than for friendship network model. The estimates show the presence of statistically significant peer effects. The higher the abilities and outcomes of student's study partners, the better is achievement of a student himself. The values of endogenous effects for help network are similar to those of a friendship network. Exogenous effects of study partners' USE in Math and their status of awardees of Olympiads are statistically insignificant.

In Table 6b the estimates of average direct, average indirect and average total impacts for study assistance network are presented.

In Table 7 the qualitative results are summarized.

The results of this study may be useful in identifying mechanisms through which peer group effects work. Our previous study (Androushchak, Poldin, Yudkevich, 2012) made on the same sample of students during their first year of study, reveals significant peer effects from the groupmates that become non-significant to their third year.. High-able students in the group had the positive influence of the students who also performed well. There were no effects from less-able groupmates. However, on the second and the third year the influence of a group in whole became insignificant. We can suppose that high-achievers can act as role models and study helpers for other students in the group. As time goes, students with close abilities form more tight friendship ties. The good student may affect the friends' attitude to learning, help them in doing homework and preparing to examinations. The influence on more distant groupmates becomes weaker. That explains why peer effect caused by the group in the second and the third year became insignificant while effect of friends and study partners is distinct.

4. Conclusion

Study of the influence of student social interactions on academic achievement is of significant interest because it helps to understand better the nature of the learning process and find the way to increase educational achievement. Correct estimation of peer effects is not an easy task. One of the difficulties is a correct identification of the students who interact during their study and therefore can influence each other. Another issue is disentangling the influence of current outcomes – that can be changed – from the effect of peers' exogenous unchangeable characteristics such as abilities.

In this work, we use data on social connections of the 3rd year students studying at one top-tier research university in Russia. Using questionnaire data, we construct network ties of two types: directed friendship ties and connections to study partners. Information about individual social ties of each student allows us to apply a spatial regression model to analysis of individual outcome. Using this model, we solve a problem of separating of overall peer effect into the influence of current achievements of peers (endogenous effect) and the influence of their abilities (exogenous effect).

We find significant positive peer effects both from academic achievement and from abilities of students. Increase of study assistants' GPA of one point is associated with the increase of student's own GPA of 0.25.

Presence of peer effects matters for educational policy. To exploit positive peer effects, the university may practice additional classes to help some students in their in studying, or provide financial aid to bright students in order to distract them from part-time working. Such help has the positive influence on the treated students and has a spillover effect to their peers. If there are endogenous peer effects, overall gain increases further due to multiplication effect. Another use of positive peer effects implies promoting teaching practice that encourages intensive social interactions, such as group project assignments.

Appendix. Interpretation of coefficients of a spatial regression model

A change in the explanatory variable for one student influence his or her own achievement (direct impact) as well as indirectly influence achievement of other students (indirect impact). The estimates of the coefficients of a spatial (network) regression take into account the interdependence between the observations (students).

Quantitative relation between dependent and independent variables is described by a marginal effect. Following LeSage and Pace (2009), we can get values for marginal effects in a network regression.

Modifying equation (4), we receive

$$\begin{aligned}\mathbf{y} &= (\mathbf{I}-\beta\mathbf{G})^{-1}\mathbf{i}\alpha + (\mathbf{I}-\beta\mathbf{G})^{-1}\mathbf{X}\boldsymbol{\gamma} + (\mathbf{I}-\beta\mathbf{G})^{-1}\mathbf{G}\mathbf{X}\boldsymbol{\delta} + (\mathbf{I}-\beta\mathbf{G})^{-1}\boldsymbol{\varepsilon} = \\ &= \mathbf{V}\,\mathbf{i}\alpha + \mathbf{V}(\mathbf{X}\boldsymbol{\gamma}+\mathbf{G}\mathbf{X}\boldsymbol{\delta}) + \mathbf{V}\,\boldsymbol{\varepsilon} = \\ &= \mathbf{V}\,\mathbf{i}\alpha + \sum_{j=1}^{k}\mathbf{S}_j\boldsymbol{x}_j + \mathbf{V}\,\boldsymbol{\varepsilon},\end{aligned} \tag{П1}$$

where

$$\begin{aligned}\mathbf{V} &= (\mathbf{I}-\beta\mathbf{G})^{-1} = \mathbf{I} + \beta\mathbf{G} + \beta^2\mathbf{G}^2 + \beta^3\mathbf{G}^3 + ..., \\ \mathbf{S}_j &= (\mathbf{I}-\beta\mathbf{G})^{-1}(\mathbf{I}\gamma_j + \mathbf{G}\delta_j).\end{aligned} \tag{П2}$$

$\boldsymbol{x}_j$ is n ×1 vector equal column j of matrix **X**, its elements are the values of exogenous variable j for n students.

From equation (П1), achievement of student i is:

$$y_i = \alpha\sum_{l=1}^{n}V_{il} + \sum_{j=1}^{k}\sum_{l=1}^{n}(S_j)_{il}\,x_{lj} + \sum_{l=1}^{n}V_{il}\varepsilon_l\,.$$

Influence of exogenous variable j on y_i consists of the sum

$$\sum_{l=1}^{n}(S_j)_{il}\,x_{lj} = (S_j)_{i1}x_{1j} + (S_j)_{i2}x_{2j} + ... + (S_j)_{in}x_{nj}\,.$$

In this equation, we can clearly see the influence on a student's outcome caused by changing the characteristics of other students. The marginal effect from the change of a continuous variable is measured by the partial derivative

$$\frac{\partial y_i}{\partial x_{lj}} = (S_j)_{il}\,. \tag{П3}$$

The marginal effect for the discrete independent variable is defined by the difference:

$$\frac{\Delta y_i}{\Delta x_{lj}} = [y_i\,|\,x_{lj}=1] - [y_i\,|\,x_{lj}=0].$$

The diagonal elements of matrix $\mathbf{S}_j$ describe direct impact on a student's achievement caused by changing his or her own characteristics. They are not equal to the parameters γ_j in equation (П2), because they take into account network interactions between students: student A affects (when $\beta\neq 0$)

the performance of other students connected to student A, and they, in turn, have a feedback effect on student A.

The non-diagonal elements of $\mathbf{S}_j$ represent indirect impact of the characteristics of student *l* on achievements of student *i*. This interdependence exists not only between friends, it is expanded further to the friends of the friends etc.

Elements of $\mathbf{S}_j$ differ and this variation reflects the difference in the position of students in a network of social ties. LeSage and Pace (2009) proposed scalar aggregated measures that measure averaged influence of exogenous variables.

Average direct impact is an average value of elements of the main diagonal of matrix $\mathbf{S}_j$:

$$ADI_j = \frac{1}{n}\sum_{l=1}^{n}(S_j)_{ll}.$$

Average total impact to an observation (student) is an average value of the sum of rows in a matrix $\mathbf{S}_j$ and is interpreted as change in achievement of a student associated with the increase of independent variable *j* of one unit for all students. Average total impact from an observation (student) measures the influence on all the students from changing a student's characteristic by one unit. These two impact measures differ by definitions, but they are equal to each other:

$$ATI_j = \frac{1}{n}\sum_{l=1}^{n}\sum_{m=1}^{n}(S_j)_{lm} = \frac{1}{n}\sum_{m=1}^{n}\sum_{l=1}^{n}(S_j)_{lm}.$$

By the definition, average indirect impact is the difference between average total impact and average direct impact:

$$AIM_j = ATI_j - ADI_j.$$

Literature

Androushchak, G.V., Poldin, O.V., and Yudkevich, M.M. (2012). Peer effects in exogenously formed university student groups. Working paper.

Arcidiacono, P., Foster, G., Goodpaster, N., and Kinsler, J. (2012). Estimating spillovers using panel data, with an application to the classroom. Quantitative Economics, 3 (3), 421-470.

Bauer, K. W., and Liang, Q. (2003). The effect of personality and precollege characteristics on first-year activities and academic performance. Journal of College Student Development, 44(3), 277-290.

Bramoullé, Y., Djebbari, H., and Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, 150, 41-55.

Brunello, G., De Paola, M., and Scoppa V. (2010). Peer effects in higher education: Does the field of study matter? *Economic Inquiry*, 48 (3), 621-634.

Burton, N. W. and Ramist, L. (2001). Predicting success in college: SAT studies of classes graduating since 1980. Research Report 2001–2. New York: College Entrance Examination Board.

Calvó-Armengol, A, Patacchini, E, and Zenou Y. (2009). Peer effects and social networks in education. *Review of Economic Studies*, 76, 1239-1267.

Carrell, S., Fullerton, R., and West J. (2009). Does your cohort matter? Measuring peer effects in college achievement. *Journal of Labor Economics*, 27 (3), 439-464.

Lavy, V., and Sand, E. (2012). The friends factor: How students' social networks affect their academic achievement and well-being. Working paper.

LeSage, J.P., and Pace, R.K. (2009). Introduction to spatial econometrics. Boca Raton, US: CRC Press Taylor & Francis Group.

Lin, X. (2010). Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics*, 28 (4), 825-860.

Lyle, D. (2007). Estimating and interpreting peer and role model effects from randomly assigned social groups at West Point. *Review of Economics and Statistics*, 89 (2), 289-299.

Lyle, D. (2009). The effects of peer group heterogeneity on the production of human capital at West Point. *American Economic Journal: Applied Economics*, 1 (4), 69-84.

Mayer, A., and Puller, S. (2008). The old boy (and girl) network: Social network formation on university campuses. Journal of Public Economics, 92, 329-347.

Noble, J. P., and Sawyer, R. L. (2002). Predicting different levels of academic success in college using high school GPA and ACT composite score. ACT Research Report Series, 1-22.

Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics*, 116, 681-704.

Shaw, E. J., Kobrin, J. L., Patterson, B. F. and Mattern K. D. (2012). The validity of the SAT for predicting cumulative grade point average by college major. Research Report 2012–6. New York: College Entrance Examination Board.

Zimmerman, D. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and Statistics*, 85 (1), 9-23.

Table 1. Descriptive statistics

Variable	Number of observations	Mean	Std. dev.	Min	Max
Friends					
USE in Russian language	167	80.0	9.6	50	100
USE in Math	167	76.4	8.0	48	100
Awardee of All-Russian Olympiad	167	0.16	0.37	0	1
Awardee of Regional Olympiad	167	0.24	0.43	0	1
GPA	167	7.35	0.88	5.57	9.21
Study assistants					
USE in Russian language	171	79.9	9.5	50	100
USE in Math	171	76.4	8.0	48	100
Awardee of All-Russian Olympiad	171	0.16	0.37	0	1
Awardee of Regional Olympiad	171	0.24	0.43	0	1
GPA	171	7.36	0.88	5.57	9.21

Table 2. Descriptive statistics of network characteristics

	Help network	Friendship network
Density	0.02	0.02
Centrality (indegree)	0.10	0.06
Reciprocity	0.19	0.42
Clasterization	0.24	0.27

Notes: Density of a network is observed proportion of ties (both in and out) in a network in relation to all the possible ties in this network. Centrality (indegree) is the observed proportion of ties going in in relation to all the possible going in ties in this network. Reciprocity is the observed proportion of mutual ties in relation to all the possible mutual ties in this network. Clasterization is a coefficient showing the probability of a tie between A and C in case when there are both connected with B.

Table 3. Estimations of peer effects for friends (without differentiation of endogenous and exogenous effects)

	Nominated ties		Reciprocal ties	
	(1)	(2)	(3)	(4)
Constant	0.075 (0.063)	0.445 (0.338)	0.239 (0.258)	0.404 (0.384)
USE in Russian language	0.023*** (4.063)	0.021*** (3.445)	0.024*** (4.306)	0.022*** (3.437)
USE in Math	0.041*** (6.097)	0.047*** (6.331)	0.036*** (5.180)	0.044*** (5.789)
Status of awardee of All-Russian Olympiad	0.808*** (5.361)		0.843*** (5.399)	
Status of awardee of Regional Olympiad	0.429*** (3.465)		0.479*** (3.838)	
Friends' mean USE in Russian language	0.002 (0.181)	-0.011 (-0.905)	-0.004 (-0.491)	-0.008 (-0.959)
Friends' mean USE in Math	0.024* (1.817)	0.032** (2.346)	0.032*** (3.222)	0.032*** (3.062)
Mean number of awardees of All-Russian Olympiad among friends	0.545** (2.213)		0.302 (1.417)	
Mean number of awardees of Regional Olympiad among friends	0.303 (1.371)		-0.073 (-0.400)	
Number of observations	167	167	156	156
R^2	0.47	0.32	0.49	0.32

Notes: *t*-statistics are in parentheses.
* significant at the 10% level.
** significant at the 5% level.
*** significant at the 1% level.

Table 4a. Estimations of peer effects for friends (with differentiation of endogenous and exogenous effects)

	Nominated ties		Reciprocal ties	
	(1)	(2)	(3)	(4)
Constant	-0.066 (-0.059)	0.140 (0.115)	-0.215 (-0.249)	-0.119 (-0.123)
USE in Russian language	0.024*** (4.497)	0.022*** (3.865)	0.025*** (4.812)	0.023*** (4.000)
USE in Math	0.038*** (5.971)	0.043*** (6.219)	0.032*** (5.004)	0.040*** (5.621)
Status of awardee of All-Russian Olympiad	0.772*** (5.443)		0.806*** (5.563)	
Status of awardee of Regional Olympiad	0.416*** (3.572)		0.448*** (3.876)	
Friends' mean USE in Russian language	-0.005 (-0.470)	-0.015 (-1.428)	-0.009 (-1.205)	-0.011 (-1.462)
Friends' mean USE in Math	0.011 (0.850)	0.012 (0.864)	0.024** (2.492)	0.018* (1.786)
Mean number of awardees of All-Russian Olympiad among friends	0.237 (0.961)		0.075 (0.363)	
Mean number of awardees of Regional Olympiad among friends	0.171 (0.802)		-0.203 (-1.183)	
β (endogenous effect)	0.253*** (2.612)	0.339*** (3.662)	0.232*** (3.172)	0.277*** (3.860)
Number of observations	167	167	156	156
R^2	0.48	0.32	0.48	0.32

Notes: *t*-statistics are in parentheses.
* significant at the 10% level.
** significant at the 5% level.
*** significant at the 1% level.

Table 4б. Average direct, indirect and total impact of exogenous characteristics for friendship network (for models (1) and (2) in Table 4a).

	(1)			(2)		
	Direct impact	Indirect impact	Total impact	Direct impact	Indirect impact	Total impact
USE in Russian language	0.024***	0.002	0.025*	0.021***	-0.011	0.010
USE in Math	0.039***	0.027	0.065***	0.044***	0.039**	0.083***
Status of awardee of All-Russian Olympiad	0.793***	0.560*	1.353***			
Status of awardee of Regional Olympiad	0.428***	0.369	0.797**			

Notes: ** significant at the 5% level, *** significant at the 1% level.

Table 5. Estimation of peer effects for groupmates.

	Without endogenous variable		With endogenous variable	
	(1)	(2)	(3)	(4)
Constant	-0.842 (-0.233)	3.164 (0.946)	-0.895 (-0.254)	2.928 (0.870)
USE in Russian language	0.025*** (4.243)	0.020*** (3.235)	0.025*** (4.377)	0.021*** (3.295)
USE in Math	0.040*** (5.720)	0.048*** (6.281)	0.040*** (5.892)	0.048*** (6.377)
Status of awardee of All-Russian Olympiad	0.906*** (6.057)		0.912*** (6.244)	
Status of awardee of Regional Olympiad	0.478*** (3.680)		0.480*** (3.799)	
Groupmates' mean USE in Russian language	0.024 (0.468)	-0.021 (-0.516)	0.029 (0.558)	-0.021 (-0.518)
Groupmates' mean USE in Math	0.008 (0.258)	0.008 (0.297)	0.015 (0.430)	0.004 (0.134)
Mean number of awardees of All-Russian Olympiad among groupmates	1.017 (1.250)		1.220 (1.306)	
Mean number of awardees of Regional Olympiad among groupmates	0.520 (0.575)		0.616 (0.673)	
β (endogenous effect)			-0.119 (-0.441)	-0.073 (-0.322)
Number of observations	167	167	167	167
R^2	0.44	0.29	0.44	0.29

Notes: *t*-statistics are in parentheses.
*** significant at the 1% level.

.

Table 6a. Estimation of peer effects for study partners.

	Without endogenous variable		With endogenous variable	
	(1)	(2)	(3)	(4)
Constant	-1.904 (-1.650)	-1.059 (-0.836)	-2.107* (-1.933)	-1.554 (-1.321)
USE in Russian language	0.024*** (4.547)	0.022*** (3.605)	0.024*** (4.756)	0.021*** (3.842)
USE in Math	0.034*** (5.210)	0.039*** (5.419)	0.032*** (5.160)	0.037*** (5.387)
Status of awardee of All-Russian Olympiad	0.857*** (6.100)		0.821*** (6.188)	
Status of awardee of Regional Olympiad	0.440*** (3.699)		0.427*** (3.801)	
Study partners' mean USE in Russian language	0.011 (1.026)	-0.004 (-0.302)	0.005 (0.464)	-0.010 (-0.919)
Study partners' mean USE in Math	0.043*** (3.969)	0.051*** (4.221)	0.031*** (2.668)	0.033*** (2.564)
Mean number of awardees of All-Russian Olympiad among study partners	0.625** (2.482)		0.380 (1.517)	
Mean number of awardees of Regional Olympiad among study partners	0.407** (1.983)		0.235 (1.156)	
β (endogenous effect)			0.250** (2.481)	0.337*** (3.520)
Number of observations	171	171	171	171
R^2	0.52	0.35	0.51	0.31

Notes: *t*-statistics are in parentheses.
** significant at the 5% level.
*** significant at the 1% level.

Table 6b. Average direct, indirect and total impact of exogenous characteristics of network of study assistants (for models (3) and (4) in Table 6a).

	(3)			(4)		
	Direct impact	Indirect impact	Total impact	Direct impact	Indirect impact	Total impact
USE in Russian language	0.026***	0.025	0.050	0.021***	-0.005	0.017
USE in Math	0.040***	0.008	0.047	0.038***	0.067***	0.106***
Status of awardee of All-Russian Olympiad	0.907***	0.947	1.854			
Status of awardee of Regional Olympiad	0.477***	0.514	0.991**			

Notes: ** significant at the 5% level, *** significant at the 1% level.

Table 7. Presence of peer group effects in the models with different specification of reference groups

	Reference group		
	Student group	Friends	Study partners
Peer effect (without differentiation of endogenous and exogenous effects)	no	yes	yes
Endogenous effect	no	yes	yes
Exogenous effect	no	no	yes