

Прикладная эконометрика, 2015, 40 (4), с. 49–63.

Applied Econometrics, 2015, 40 (4), pp. 49–63.

Е. В. Креховец, О. В. Польшин<sup>1</sup>

## Факторы формирования дружеских связей студентов

*В работе рассматривается сетевая структура межличностных коммуникативных (дружеских) связей студентов университета и анализируются факторы, влияющие на формирование таких сетей. С использованием данных опроса студентов и административной информации проводится эконометрическое оценивание моделей образования дружеских связей, основанных на гипотезах гомофилии и эффекта соседства. В частности, оцениваются линейная и логистическая модели с фиксированными эффектами и без фиксированных эффектов. Полученные результаты свидетельствуют о статистически значимой связи между вероятностью возникновения дружбы и такими факторами, как принадлежность к одной студенческой группе, проживание в общежитии, совпадение пола и близость успеваемости. При увеличении числа совпадающих факторов вероятность связи растет.*

**Ключевые слова:** социальные сети; сети дружбы; социальный капитал; высшее образование; модель бинарного выбора.

**JEL classification:** I21; Z13.

### 1. Введение

Наряду с традиционными образовательными функциями, институт высшего образования играет важную роль в формировании социального капитала учащихся. Социальный капитал представляет собой ресурс, который люди получают через особые социальные структуры и в дальнейшем применяют в своих целях (Baker, 1990). В отличие от других форм капитала, социальный капитал принадлежит не отдельному человеку, а социальной структуре, в которой он находится, но в то же время такой капитал может использоваться каждым человеком для облегчения производственной деятельности и достижения индивидуальных целей. Социальный капитал может быть представлен как сеть социальных связей разного уровня (Радаев, 2002). Такое определение социального капитала позволяет использовать методологию и инструментарий анализа социальных сетей для исследования формирования социального капитала учащихся.

В настоящей работе исследуется формирование социального капитала студентов в академической среде с помощью анализа студенческих социальных сетей. В работе (Burt, 1992) отмечается, что социальный капитал — это дружеские, рабочие и более общие контакты, с помощью которых индивид получает возможность использовать свой финансовый и че-

<sup>1</sup> Креховец Екатерина Владимировна — НИУ ВШЭ, Нижний Новгород; [ekrehovets@hse.ru](mailto:ekrehovets@hse.ru).  
Польшин Олег Викторович — НИУ ВШЭ, Москва; [opoldin@hse.ru](mailto:opoldin@hse.ru).

ловеческий капитал. В данной работе в качестве социального капитала рассматриваются дружеские связи студентов. Для студента дружеские связи являются не только неформальными отношениями в рамках института высшего образования, но и весьма полезным ресурсом, как во время студенческой жизни, так и после окончания университета, например, при поиске работы. Дружеские социальные связи могут использоваться учащимся для улучшения своей академической успеваемости. Данный эффект известен в литературе как эффект сообучения и может рассматриваться как пример формирования человеческого капитала за счет ресурсов социального капитала индивида. После окончания университета созданные в нем социальные связи можно применить для построения карьеры. Обладая широкими социальными связями и входя в определенный социальный круг, студент может получить более выгодную позицию при трудоустройстве. Таким образом, социальный капитал можно представить как ресурс для создания финансового капитала. Далее в работе социальный капитал студентов будет определяться как социальная сеть.

Анализируя социальную сеть в качестве зависимой переменной, исследователи выделяют два основных механизма ее формирования: гомофилия и эффект соседства. Гомофилия в данном случае означает, что студенты предпочитают дружить со сверстниками из собственной этнической и расовой группы со схожими социально-экономическими характеристиками и интересами. Выбор одинаковых программ обучения, равно как и участие в одних и тех же университетских обществах, увеличивает вероятность возникновения дружеских связей между студентами. Кроме перечисленных выше факторов, имеет значение и географическая удаленность — чем ближе студенты живут друг к другу, тем вероятней они будут находиться в одной социальной сети (эффект соседства).

Гипотеза гомофилии не является единственной, можно выдвинуть аргументы и противоположного толка. Так, полезность социальных сетевых связей во многом обусловлена тем, что эти связи обеспечивают взаимодействие между людьми с различными мыслями, идеями и знаниями (Granovetter, 1973; Vurt, 1992). Поэтому студенты могут предпочитать не только схожих с ними однокурсников, стремясь к разнообразию круга общения.

Исследований факторов социальных связей среди российских студентов очень мало, и настоящая работа в какой-то мере восполняет этот пробел. В статье на примере студентов первого курса одного российского университета оцениваются эконометрические модели вероятности формирования дружеских связей между двумя студентами. В качестве факторов используются легко наблюдаемые характеристики. Выборка студентов сформирована из административных данных и результатов опроса студентов четырех факультетов российского вуза.

## 2. Обзор литературы

В литературе, посвященной изучению студенческих социальных сетей, выделяется несколько направлений. Во-первых, следует отметить работы, посвященные теоретическому моделированию сетей учащихся и оценке их характеристик, во-вторых, это работы, анализирующие влияние социальных сетей на академические успехи учащихся (эффекты среды или сообучения), и, наконец, эмпирические работы по изучению формирования социальных сетей, оценке наиболее значимых факторов, определяющих возникновение связей в сети. Ниже будет представлен обзор некоторых работ, посвященных вопросу формирования

социальных связей среди учащихся, в которых определяются ключевые факторы, оказывающие воздействие на возникновение и структуру социальных сетей.

Механизм образования связей по принципу близости индивидуальных черт (гомофилия) характерен для различных типов социальных сетей. В результате личные социальные контакты людей происходят в значительной степени с людьми, близкими по социально-демографическим и личностным характеристикам (McPherson et al., 2001). В случае студенческих сетей гомофилия проявляется в склонности учащихся дружить с близкими по различным характеристикам сверстниками.

Многие зарубежные исследователи обращаются к вопросу о влиянии этничности и расовой принадлежности на формирование дружеских связей среди студентов. Сравнивая группу американских и европейских студентов с группой студентов из Азии, Латинской Америки и Африки, Kenny и Stryker (1994, 1996) пришли к выводу о том, что социальные сети формируются, среди прочего, под влиянием расовой и этнической принадлежности. Данный эффект известен как «расовая гомофилия». Было обнаружено, что американские и европейские студенты создают более обширные и густые сети в отличие от своих однокурсников, принадлежащих другим этническим группам. Также было отмечено, что адаптация студентов в колледже происходит по-разному — студенты европеоидной расы получают большую поддержку от своих друзей, в то время как их сверстники из других этнических групп получают большую поддержку не от друзей, а от родственников и семьи.

Подобный эффект расовой гомофилии был найден и в исследованиях (Mayer, Puller, 2008), основанных на данных социальной сети Facebook в США. В этой работе расовое единообразие определяется как доминирующий фактор в процессе формирования дружеских связей учащихся, характерный для разного типа учебных заведений, от частных закрытых школ до крупных государственных университетов. Отмечается, впрочем, что политика университета, направленная на увеличение контактов между представителями разных этнических и расовых групп, может положительно повлиять на формирование дружеских социальных сетей среди этих студентов. В работе (Wejnert, 2010) также тестируется и находит подтверждение гипотеза о возможностях университетской среды влиять на формирование межнациональных и межрасовых связей среди студентов. Wimmer и Lewis (2010) также обратили внимание на существенное влияние расового подобию на создание дружеской сети студентов. Авторы отмечают роль близости по другим характеристикам, в частности, по социально-демографическому статусу.

Моделируя формирование студенческих социальных сетей, Marmaros и Sacerdote (2006) получили результаты, согласно которым ключевыми факторами, определяющими возникновение социальных сетей в студенческой среде, являются принадлежность к одной этнической группе или расе, географическая близость, семейное окружение, а также общие интересы в рамках обучения в университете (такие как изучаемые курсы или участие в университетских спортивных секциях). Один из основных результатов данного исследования свидетельствует о том, что вероятность возникновения дружбы между студентами одной расовой группы значительно выше, чем между студентами из разных расовых групп. Однако авторы полагают, что совместное обучение со сверстниками из разных этнических групп, особенно на первом году обучения, увеличивает шансы на формирование интернациональных дружеских связей.

В российских вузах вопросы расы и этничности не столь злободневны, как в США. Тем не менее, после введения ЕГЭ в ведущих вузах возросло число студентов из других реги-

онов, ежегодно увеличивается поток иностранных студентов, в том числе из стран СНГ и Балтии (Гаврилова, Яценко, 2012).

Наблюдается также существование гомофилии по другим характеристикам, таким как социально-экономический статус учащихся, схожесть хобби и интересов, выбор одинаковых «учебных траекторий». В исследовании (Lee et al., 2011) по данным немецкой социальной сети StudyVZ отмечается, что студенты технических и математических направлений с большей вероятностью формируют дружеские связи со студентами схожих направлений, чем со студентами гуманитарных и художественных университетов. В другом исследовании найдено, что для студентов первого курса детерминантами связей являются схожесть по половому и расовому признаку, в то время как для более старших курсов возрастает значимость общих интересов, например, религиозные убеждения, членство в клубах (Godley, 2008).

В работе (Elias et al., 2007), анализирующей данные по учащимся Аргентины, делаются выводы о значимости таких факторов, как успеваемость студентов и внешняя привлекательность. Согласно исследованию, вероятность сформировать дружеские связи у внешне привлекательного учащегося с высокими академическими достижениями выше, чем у тех, кто не обладает подобными качествами. Авторы также тестируют модель для школ как с совместным, так и с раздельным обучением. В результате делается заключение о значимости академических успехов для обоих типов школ, в то время как внешняя привлекательность оказывается важным фактором только для смешанных школ.

Исследование (Soetevent, Koogeman, 2005) также приводит доказательства влияния факторов сходства на вероятность возникновения дружеских связей. Значимость отдельных параметров варьируется в зависимости от типа поведения студентов. Так, например, в случае личных встреч большее влияние оказывают факторы этнического и расового характера, тогда как при онлайн коммуникации или телефонных звонках более значимым оказывается фактор гендерной принадлежности.

Кроме определения механизма формирования дружеских связей среди студентов за счет детерминантов, связанных с единообразием и подобием характеристик учащихся, в литературе находятся убедительные свидетельства влияния факторов географической близости студентов, называемые «эффект близости» или «эффект соседства». Эффект близости (или соседства) связан с совместным проживанием студентов в одной комнате в общежитии и географической близостью между местами их проживания и обучения (Festinger et al., 1950).

Анализируя данные германской социальной сети StudyVZ, исследователи пришли к выводу о непосредственном влиянии географического местоположения на формирование дружеских связей среди студентов разных вузов Германии (Lee et al., 2011). Согласно полученным результатам, удаленность, соответствующая 100 минутам езды на автомобиле, уменьшает вероятность формирования дружеских связей на 91%, дружеские связи чаще формируются среди студентов, чьи университеты находятся в одной федеральной области.

Аналогичные выводы относительно значимости географического расстояния в процессе формирования дружеских связей получены в (Preciado et al., 2012). Исследователи выявили зависимость между вероятностью возникновения дружбы и географической удаленностью двух учащихся. Кроме того, для учеников из разных учебных заведений такая зависимость более сильная, чем для тех, кто обучается в одной школе.

В другой работе (Traud et al., 2011), анализирующей данные социальной сети Facebook, также подтвердилось предположение об эффекте близости при формировании дружеских

связей. Однако было показано, что не существует одного доминирующего фактора, и только взаимодействие нескольких характеристик может объяснить процесс формирования студенческой социальной сети.

В исследовании Foster (2005) также делается вывод о том, что процесс формирования дружеских связей среди студентов зависит от совокупности факторов, относящихся как к эффектам гомофилии (таких как раса, пол, успеваемость), так и к эффектам географической близости (например, обучение в одном кампусе).

В работе (Валеева и др., 2013) изучались факторы образования связей между студентами одного из факультетов российского вуза методом экспоненциального моделирования случайных графов (exponential random graph model). Было показано, что связи дружбы образуются чаще между студентами одного пола, обучающимися в одной группе. Чем меньше разница в успеваемости между двумя студентами, тем больше вероятность образования между ними связей.

В отличие от (Валеева и др., 2013), настоящая работа использует данные о значительно большем числе студентов, представляющих четыре факультета, что позволяет оценить, как варьируются результаты от факультета к факультету. Также вместо моделирования сети в целом оценивается эконометрическая модель для диадных (парных) связей. Результаты оценивания такой модели проще интерпретируются и могут быть непосредственно использованы, например, для прогнозирования связи между двумя случайно выбранными студентами.

### 3. Данные

В рамках данного исследования был проведен опрос студентов первого курса четырех факультетов НИУ ВШЭ – Нижний Новгород: бизнес-информатики, менеджмента, экономики, права. При проведении опроса респондентам было предложено указать свою фамилию (или номер студенческого билета). Вопрос о социальных связях был сформулирован следующим образом: «Напишите, пожалуйста, имена и фамилии своих однокурсников, с которыми Вы общаетесь больше всего (можно указать неограниченное число однокурсников)». Таких однокурсников далее будем называть друзьями студента, который их номинировал. На момент опроса студенты не знали, отметил ли кто-то их самих в качестве друзей. На факультете бизнес-информатики было составлено 403 дружеских связей, на факультете менеджмента — 456, экономики — 494, права — 192. Описательная статистика данных представлена в табл. 1.

Средний балл брался из рейтингов студентов, публикуемых деканатами соответствующих факультетов. Публикация рейтингов происходила после проведения опроса, т.е. респонденты на момент опроса не знали о величине среднего балла своих однокурсников. На факультетах экономики и менеджмента большую долю среди студентов составляют девушки. На факультете бизнес-информатики доля мужчин чуть выше, чем доля девушек. Процент учащихся, проживающих в общежитии, определялся на основании ответов респондентов на вопрос: «Где Вы проживаете на данный момент?». В зависимости от ответа на этот вопрос, множество студентов разбивается на две категории: тех, кто проживает в общежитии, и тех, кто не живет в общежитии (живет с родственниками, в своем или арендуемом жилье).

Таблица 1. Описательная статистика выборки

	Число наблюдений	Среднее	Стандартное отклонение	Минимум	Максимум
<i>Факультет бизнес-информатики</i>					
Доля мужчин, %	81	54.3			
Доля проживающих в общежитии, %	81	9.9			
Средний балл	81	6.74	1.07	5.00	9.32
<i>Факультет менеджмента</i>					
Доля мужчин, %	87	27.6			
Доля проживающих в общежитии, %	87	21.8			
Средний балл	87	7.22	0.86	4.65	9.15
<i>Факультет экономики</i>					
Доля мужчин, %	98	25.5			
Доля проживающих в общежитии, %	98	16.3			
Средний балл	98	7.11	1.05	5.25	9.22
<i>Факультет права</i>					
Доля мужчин, %	45	33.3			
Доля проживающих в общежитии, %	45	15.6			
Средний балл	45	7.22	1.03	4.90	9.35

Примечание. Максимально возможное значение среднего балла равно 10.

#### 4. Методология и результаты оценивания

Предполагается, что каждый студент может формировать дружеские связи с любым другим однокурсником своего факультета. Каждое наблюдение относится к паре, составленной из двух студентов одного факультета. Число наблюдений, таким образом, соответствует числу всевозможных пар.

Зависимая бинарная переменная описывает статус дружеской связи:  $F_{ij} = 1$ , если студент  $i$  назвал студента  $j$  своим другом,  $F_{ij} = 0$  — если нет. Связь может быть не взаимной. Вероятность дружбы предполагается зависящей от индивидуальных характеристик каждого студента и характеристик пары.

Зависимость от индивидуальных характеристик будем учитывать с помощью фиксированных эффектов  $\theta^i$  и  $\theta^j$ . Это позволяет учесть не только наблюдаемые, но и ненаблюдаемые характеристики студентов и избежать смещения оценок вследствие пропущенных переменных. Что касается зависимости от характеристик пары, авторы опираются на гипотезу гомофилии в образовании связей и предполагают дружбу функцией от близости наблюдаемых индивидуальных характеристик. Для описания близости таких дискретных индивидуальных характеристик, как пол, принадлежность к конкретной студенческой группе, проживание в общежитии, вводятся бинарные переменные-индикаторы, принимающие значение 1 в случае совпадения этих характеристик у рассматриваемой пары студентов, и значение 0 в случае, если эти характеристики различаются:

$$d(x^{ik}, x^{jk}) = I(x^{ik} = x^{jk}). \quad (1)$$

Для среднего балла степень близости описывается через абсолютные значения разности индивидуальных значений:

$$d(x^{ik}, x^{jk}) = -|x^{ik} - x^{jk}|, \quad (2)$$

причем средний балл по факультетам предварительно стандартизирован.

С учетом введенных обозначений, оцениваемая эконометрическая модель имеет вид:

$$p_t = \Pr(F_{ij} = 1) = f\left(\beta_0 + \sum_{k=1}^K \beta_k d(x_t^{ik}, x_t^{jk}) + \theta^i + \theta^j\right) = f(d_t' \beta + \theta^i + \theta^j), \quad t = 1, \dots, N, \quad (3)$$

где  $p_t$  — вероятность дружеской связи,  $d(x^{ik}, x^{jk})$  — функция близости  $k$ -й характеристики у студентов  $i$  и  $j$ ,  $\theta^i$  и  $\theta^j$  — фиксированные эффекты,  $N$  — число пар студентов,  $K$  — число характеристик у студентов.

В качестве функции  $f(\cdot)$  в правой части выражения (3) используются линейная и логистическая функции. Известный недостаток линейной модели состоит в том, что при некоторых значениях объясняющих переменных и некоторых значениях параметров предсказываемая вероятность выходит за пределы интервала  $[0, 1]$  (Магнус и др., 2007). Рассмотрение линейной модели для бинарной зависимой переменной обусловлено следующими соображениями. При оценивании нелинейных моделей с фиксированными эффектами число фиктивных переменных растет с увеличением числа индивидуумов в выборке, что может привести к смещению и несостоятельности оценок (incidental parameters problem), см., например, (Greene, 2012; Wooldridge, 2010). В нашем случае проблема не столь выражена, как при оценивании панельных данных с небольшим числом временных интервалов, поскольку число наблюдений, приходящихся на одного студента, растет при увеличении социальной сети. Тем не менее, наряду с нелинейной логистической моделью, оцениваются и линейные регрессии<sup>2</sup>.

Так как анализируются данные из всевозможных комбинаций пар студентов, то данные об одном и том же студенте появляются во множестве пар, и различные наблюдения не являются полностью независимыми. Помимо использования фиксированных эффектов, существует альтернативная процедура (multiple regression quadratic assignment procedure, MRQAP), смысл которой заключается в тестировании нулевой гипотезы относительно регрессионных оценок с помощью перестановок (Dekker et al., 2007; Krackhardt, 1988). Идею метода можно описать следующим образом. Представим анализируемые переменные в матричном виде так, что номер строки соответствует первому студенту в паре, а номер столбца — второму студенту. На первом шаге стандартным методом наименьших квадратов оценивается зависимость вероятности дружеской связи от характеристик студенческой пары  $i, j$ :

$$p^{ij} = f\left(\beta_0 + \sum_{k=1}^K \beta_k d(x^{ik}, x^{jk})\right). \quad (4)$$

<sup>2</sup> Некоторые авторы считают, что линейные модели дают хорошие оценки предельных эффектов (Wooldridge, 2010; Angrist, Pischke, 2008).

Затем в матрице зависимой переменной переставляются случайным образом строки и столбцы, причем перестановка строк должна совпадать с перестановкой столбцов (чтобы данные по строкам и столбцам продолжали относиться к одному человеку), после чего снова делается оценивание по МНК. Поскольку после перестановки друзья и не друзья перемешиваются, в соответствии с нулевой гипотезой ожидается отсутствие связи между зависимой и независимой переменными. Повторяя перемешивание и оценивание много раз, можно получить эмпирическое распределение коэффициентов регрессии при нулевой гипотезе. Маржинальный уровень значимости коэффициентов ( $p$ -значение) находится как доля оценок после перемешивания, модуль оценки которых оказывается не ниже оценки для исходных данных.

Оценки моделей вероятности возникновения дружбы между студентами представлены в табл. 2 для факультетов бизнес-информатики, менеджмента, экономики и права<sup>3</sup>. Число наблюдений соответствует числу пар студентов. Столбец 1 содержит оценки модели (4), где  $p$ -значения рассчитаны MRQAP процедурой по 1000 перестановок. В столбце 2 приведены оценки модели с фиксированными эффектами (FE), заданной уравнением (3). В столбце 3 приведены оценки логистической модели,  $p$ -значения рассчитаны MRQAP алгоритмом. В столбце 4 показаны оценки логистической модели с фиксированными эффектами.

Для логистических моделей рассчитаны предельные эффекты, которые можно сравнивать с оценками линейных спецификаций. В столбцах 5 и 6 показаны значения средних предельных эффектов от изменения соответствующей зависимой переменной в логистической модели без фиксированных эффектов и с фиксированными эффектами<sup>4</sup>. Для непрерывной переменной (близость успеваемости) средний предельный эффект рассчитывается как среднее выборочное значение:

$$AME_k = \frac{1}{N} \sum_{i=1}^N \frac{\partial p_i(d'_i \hat{\beta})}{\partial d_k} = \frac{1}{N} \sum_{i=1}^N \frac{e^{d'_i \hat{\beta}}}{(1 + e^{d'_i \hat{\beta}})^2} \hat{\beta}_k. \quad (5)$$

При расчете среднего предельного эффекта для дискретных переменных производная в (5) заменяется разностью:

$$AME_k = \frac{1}{N} \sum_{i=1}^N \left( p_i(d'_i \hat{\beta} | d_k = 1) - p_i(d'_i \hat{\beta} | d_k = 0) \right). \quad (6)$$

<sup>3</sup> Оценки в столбцах 1 и 2 получены МНК, в столбцах 3–6 — методом максимального правдоподобия. Для столбцов 1 и 3  $p$ -значения рассчитаны по методу MRQAP. Расчеты производились с использованием статистического пакета STATA, MRQAP процедура проводилась в среде R (команда *nlogit* пакета *sna*).

<sup>4</sup> Число наблюдений при оценивании логистической модели с фиксированными эффектами (столбцы 4 и 6) меньше, чем в других. Это вызвано тем, что некоторые студенты не номинировали других студентов из анализируемой выборки, и поэтому соответствующие им фиксированные эффекты безошибочно предсказывают зависимую переменную, делая невозможной оценку модели. Такие наблюдения автоматически исключались из выборки.

**Таблица 2.** Оценки линейной и логистической моделей вероятности возникновения дружбы на различных факультетах

Независимые переменные	Линейная модель		Логистическая модель		Предельные эффекты для логистической модели	
	MRQAP	FE	MRQAP	FE		FE
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Факультет бизнес-информатики</i>						
Обучение в одной группе	0.275*** (<0.01)	0.280*** (<0.01)	4.104*** (<0.01)	4.525*** (<0.01)	0.179*** (<0.01)	0.182*** (<0.01)
Одинаковый пол	0.030*** (<0.01)	0.029*** (<0.01)	0.707*** (<0.01)	1.014*** (<0.01)	0.031*** (<0.01)	0.041*** (<0.01)
Оба живут в общежитии/ не в общежитии	0.026*** (0.004)	0.029* (0.055)	0.641*** (0.002)	0.806*** (<0.01)	0.028*** (<0.01)	0.033*** (0.009)
Близость среднего балла	0.019*** (<0.01)	0.023*** (<0.01)	0.513*** (<0.01)	0.616*** (<0.01)	0.022*** (<0.01)	0.025*** (<0.01)
Константа	-0.007 (0.461)	-0.019 (0.557)	-5.434*** (<0.01)	-6.486*** (<0.01)	—	
Число наблюдений	6480	6480	6480	6086	6480	6086
$R^2$	0.219	0.244				
Псевдо $R^2$			0.376	0.445	0.376	0.445
<i>Факультет менеджмента</i>						
Обучение в одной группе	0.198*** (<0.01)	0.196*** (<0.01)	3.072*** (<0.01)	3.389*** (<0.01)	0.147*** (<0.01)	0.148*** (<0.01)
Одинаковый пол	0.032*** (<0.01)	0.060*** (<0.01)	0.704*** (<0.01)	1.341*** (<0.01)	0.034*** (<0.01)	0.059*** (<0.01)
Оба живут в общежитии/ не в общежитии	0.030*** (<0.01)	0.054*** (<0.01)	0.574*** (0.001)	1.130*** (<0.01)	0.028*** (<0.01)	0.049*** (<0.01)
Близость среднего балла	0.019*** (<0.01)	0.0360*** (<0.01)	0.430*** (<0.01)	0.771*** (<0.01)	0.021*** (<0.01)	0.034*** (<0.01)
Константа	-0.007 (0.395)	0.018 (0.622)	-4.870*** (<0.01)	-4.841*** (<0.01)		
Число наблюдений	7482	7482	7482	7226	7482	7226
$R^2$	0.140	0.174				
Псевдо $R^2$			0.263	0.353	0.263	0.353
<i>Факультет экономики</i>						
Обучение в одной группе	0.149*** (<0.01)	0.154*** (<0.01)	2.304*** (<0.01)	2.521*** (<0.01)	0.101*** (<0.01)	0.110*** (<0.01)
Одинаковый пол	0.038*** (<0.01)	0.036*** (<0.01)	0.936*** (<0.01)	0.932*** (<0.01)	0.041*** (<0.01)	0.041*** (<0.01)
Оба живут в общежитии/ не в общежитии	0.008 (0.267)	0.016* (0.084)	0.159 (0.305)	0.375** (0.035)	0.007 (0.153)	0.016** (0.036)
Близость среднего балла	0.027*** (<0.01)	0.032*** (<0.01)	0.647*** (<0.01)	0.718*** (<0.01)	0.028*** (<0.01)	0.031*** (<0.01)

Окончание табл. 2

Независимые переменные	Линейная модель		Логистическая модель		Предельные эффекты для логистической модели	
	MRQAP	FE	MRQAP	FE	FE	
	(1)	(2)	(3)	(4)	(5)	(6)
Константа	0.024** (0.013)	-0.042 (0.115)	-3.990*** (<0.01)	-5.490*** (<0.01)		
Число наблюдений	9506	9506	9506	8832	9506	8832
$R^2$	0.091	0.117				
Псевдо $R^2$			0.188	0.251	0.188	0.251
<i>Факультет права</i>						
Обучение в одной группе	0.201*** (<0.01)	0.198*** (<0.01)	2.254*** (<0.01)	2.596*** (<0.01)	0.172*** (<0.01)	0.192*** (<0.01)
Одинаковый пол	0.063*** (<0.01)	0.074*** (<0.01)	0.871*** (<0.01)	0.975*** (<0.01)	0.066*** (<0.01)	0.072*** (<0.01)
Оба живут в общежитии/ не в общежитии	0.031 (0.131)	0.062** (0.034)	0.421 (0.138)	0.705** (0.025)	0.032** (0.025)	0.052** (0.026)
Близость среднего балла	0.014 (0.198)	0.046*** (<0.01)	0.198 (0.171)	0.694*** (<0.01)	0.015* (0.069)	0.051*** (<0.01)
Константа	-0.009 (0.706)	-0.128** (0.048)	-4.097*** (<0.01)	-6.091*** (<0.01)		
Число наблюдений	1980	1980	1980	1684	1980	1684
$R^2$	0.118	0.200				
Псевдо $R^2$			0.179	0.303	0.179	0.303

Примечание. В скобках указаны  $p$ -значения. \*, \*\*, \*\*\* — значимость на уровне 10, 5 и 1% соответственно.

Сравнение оценок в столбцах 1, 2, 5, 6 позволяет сделать вывод о том, что различные методы в большинстве случаев дают качественно сопоставимые результаты. Результаты свидетельствуют о том, что наиболее сильным по влиянию оказался фактор соседства (переменные «оба живут в общежитии / не в общежитии» и «обучение в одной группе»). Обучение в одной группе увеличивает вероятность возникновения дружбы, в зависимости от спецификации модели и факультета, от 10 до 28%. Роль места проживания гораздо слабее — от 1 до 6%, причем оценки не всегда значимы. Полученные результаты позволяют сделать вывод о существенной роли географической близости при формировании дружеских взаимодействий в социальных сетях студентов.

Совпадение пола студентов статистически значимо сопряжено с ростом вероятности возникновения дружбы на 3–7%. Близость успеваемости — другая значимая характеристика. Сокращение разницы среднего балла на одно стандартное отклонение увеличивает вероятность возникновения дружбы между студентами на 1.5–5%.

Интересным является вопрос о том, как вероятность дружбы ведет себя при возрастании числа совпадающих характеристик студентов. Чтобы ответить на него, оценивалась логистическая модель с фиксированными эффектами:

$$p_i = \Pr(F_{ij} = 1) = f\left(\beta_0 + \sum_{k=1}^{15} \beta_k D_k + \theta^i + \theta^j\right), \quad (7)$$

$$D_k = d_1 \times d_2 \times d_3 \times d_4, \quad d_n = 0, 1, \quad n = 1, 2, 3, 4.$$

В этой модели зависимой переменной является вероятность связи между парой студентов, а независимыми переменными — всевозможные комбинации бинарных переменных, описывающих близость характеристик этой пары: Например, при совпадении студенческой группы у двух студентов соответствующая переменная ( $d_1$ ) полагалась равной единице, при несовпадении группы — нулю. Аналогично строились переменные для пола ( $d_2$ ), проживания в общежитии ( $d_3$ ) и близости среднего балла ( $d_4$ )<sup>5</sup>. Во избежание полной мультиколлинеарности была исключена комбинация из четырех нулей, соответствующая несовпадению ни одного из четырех признаков.

По оцененной модели были рассчитаны вероятности для данных комбинаций<sup>6</sup>. В таблице 3 представлены результаты оценивания, упорядоченные в порядке возрастания вероятности. Расчеты выполнены для выборки, объединяющей все факультеты. Как видно из таблицы, увеличение числа близких характеристик увеличивает вероятность дружеской связи. При совпадении всех четырех атрибутов вероятность дружбы приближается к 43%, что на два порядка превышает вероятность дружбы между двумя студентами, не имеющими ни одной общей черты. Таким образом, в социальных сетях наблюдается тенденция к формированию связей между студентами, близкими по характеристикам.

## 5. Заключение

В данной работе анализировалось, как некоторые легко наблюдаемые факторы связаны с формированием дружеских социальных связей в студенческой среде. Наиболее сильно увеличивает вероятность возникновения дружеских связей принадлежность студентов к одной группе<sup>7</sup>. Совпадение условий проживания (в общежитии или нет) также связано с формированием связей. Дружеские отношения ощутимо чаще возникают среди студентов одного пола. Кроме того, значимым оказывается близость среднего балла: студенты с близкой успеваемостью будут дружить с большей вероятностью, чем студенты, чьи средние баллы сильно отличаются. Полученные качественные результаты являются устойчивыми. Они повторяются для сетей четырех различных факультетов и различных спецификаций модели — линейной вероятностной модели с фиксированными эффектами и без фиксированных эффектов, а также логистической модели.

Количественная оценка роли разных факторов важна для понимания механизма социальных связей. Круг общения студента важен для его успеваемости: друзья участвуют в формировании отношения к учебе и своими успехами способны стимулировать собственное

<sup>5</sup> Для удобства переменная «близость среднего балла» также была дискретизирована: значения функции (2), которые меньше медианы, полагались равными 1; значения выше медианы — 0.

<sup>6</sup> В пакете STATA (версия 12) эта процедура реализуется командой *margins*.

<sup>7</sup> Административно формируемые студенческие группы не являются повсеместно распространенным явлением, но характерны для учебного процесса в вузах России и некоторых других стран.

**Таблица 3.** Зависимость вероятности возникновения дружбы от сочетания факторов. Логистическая модель с фиксированными эффектами

Группа	Совпадающие факторы			Вероятность	95%-ный доверительный интервал	
	Пол	Проживание	Средний балл		От	До
0	0	0	0	0.004	0.001	0.008
0	0	0	1	0.011	0.005	0.016
0	0	1	0	0.011	0.007	0.015
0	1	0	0	0.012	0.007	0.017
0	0	1	1	0.015	0.010	0.020
0	1	0	1	0.017	0.011	0.022
0	1	1	0	0.020	0.015	0.025
0	1	1	1	0.029	0.023	0.035
1	0	0	0	0.035	0.017	0.053
1	0	0	1	0.087	0.058	0.117
1	1	0	0	0.136	0.103	0.168
1	0	1	0	0.146	0.121	0.170
1	0	1	1	0.204	0.173	0.234
1	1	1	0	0.267	0.239	0.296
1	1	0	1	0.294	0.248	0.339
1	1	1	1	0.427	0.397	0.457

*Примечание.* Вероятности и доверительные интервалы для вероятностей рассчитаны при средних значениях индивидуальных эффектов.

усердие студента, а также оказывать ему прямую поддержку, объясняя учебный материал и помогая в выполнении заданий. Предпочтение студентов контактировать с однокурсниками, близкими по социально-демографическим и личностным характеристикам, накладывает ограничения на получаемую от них информацию. Хотя выбор друзей, несомненно, является добровольным, на вероятность его формирования сильно влияют такие контролируемые факторы, как студенческая группа и общежитие. Поэтому администрация вуза может целенаправленно формировать студенческие группы и расселять студентов в общежитии, например, чтобы оптимизировать общую успеваемость, или создать благоприятную среду для самых способных студентов, либо поддержать недостаточно подготовленных учащихся.

**Благодарности.** Работа выполнена в рамках проекта «Эмпирические методы в сравнительном институциональном анализе» Программы фундаментальных исследований НИУ ВШЭ в 2015 г. и при поддержке Правительства РФ в рамках реализации «Дорожной карты» Программы 5/100 НИУ ВШЭ.

Авторы благодарны анонимным рецензентам за замечания, которые позволили улучшить статью.

### Список литературы

Валева Д. Р., Польдин О. В., Юдкевич М. М. (2013). Связи дружбы и помощи при обучении в университете. *Вопросы образования*, 4, 70–81.

- Гаврилова А. К., Яценко Е. Б. (2012). *Учебная миграция из стран СНГ и Балтии: потенциал и перспективы для России*. М.: Фонд «Наследие Евразии».
- Магнус Я. Р., Катышев П. К., Пересецкий А. А. (2007). *Эконометрика. Начальный курс*. М.: Дело.
- Радаев В. В. (2002). Понятие капитала, формы капиталов и их конвертация. *Экономическая социология*, 3 (4), 20–32.
- Angrist J., Pischke S. (2008). *Mostly harmless econometrics: An empiricists' companion*. Princeton: Princeton University Press.
- Baker W. (1990). Market networks and corporate behavior. *The American Journal of Sociology*, 96 (3), 589–625.
- Burt R. S. (1992). *Structural holes*. Cambridge: Harvard University Press.
- Dekker D., Krackhardt D., Snijders T. A. B. (2007). Sensitivity of MRQAP tests to collinearity and autocorrelation conditions. *Psychometrika*, 72 (4), 563–581.
- Elias J., Elias V., Ronconi L. (2007). Discrimination and social networks: Popularity among high school students in Argentina. *Inter-American Development Bank Research Network Working Paper R-539*.
- Godley J. (2008). Preference or propinquity? The relative contribution of selection and opportunity to friendship homophily in college. *Connections*, 1, 65–80.
- Granovetter M. (1973). The strength of weak ties. *American Journal of Sociology*, 78 (6), 1360–1380.
- Greene W. (2012). *Econometric analysis. 7<sup>th</sup> Edition*. Prentice Hall.
- Festinger L., Schachter S., Back K. (1950). *Social pressure in informal groups*. Stanford: Stanford University Press.
- Foster G. (2005). Making friends: A nonexperimental analysis of social pair formation. *Human Relations*, 58, 1443–1465.
- 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: *102<sup>nd</sup> annual meeting of the American Psychological Association*, Los Angeles, CA, USA.
- Kenny M., Stryker S. (1996). Social network characteristics and college adjustment among racially and ethnically diverse first-year students. *Journal of College Student Development*, 37 (6), 649–658.
- Krackhardt D. (1988). Predicting with networks: Nonparametric multiple-regression analysis of dyadic data. *Social Networks*, 10 (4), 359–381.
- 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 friendship form? *The Quarterly Journal of Economics*, 121 (1), 79–119.
- McPherson M., Smith-Lovin L., Cook J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27, 415–444.
- Mayer A., Puller S. L. (2008). The old boy (and girl) network: Social network formation on university campuses. *Journal of Public Economics*, 92, 329–347.
- Preciado P., Snijders T., Burk W. J., Stattin H., Kerr M. (2012). Does proximity matter? Distance dependence of adolescent friendships. *Social Networks*, 34, 18–31.
- Soetevent A. R., Kooreman P. (2005). Social ties within school classes: The roles of gender, ethnicity, and having older siblings. *Oxford Review of Economic Policy*, 21, 373–391.
- 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.

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.

Wooldridge J. (2010). *Econometric analysis of cross section and panel data. 2nd Edition*. MIT Press.

Поступила в редакцию 16.03.2015;  
принята в печать 22.05.2015.

---

Krekhovets E., Poldin O. An empirical analysis of students' friendship ties formation. *Applied Econometrics*, 2015, 40 (4), pp. 49–63.

---

### **Ekaterina Krekhovets**

National Research University Higher School of Economics, Nizhny Novgorod, Russian Federation;  
ekrekhovets@hse.ru

### **Oleg Poldin**

National Research University Higher School of Economics, Moscow, Russian Federation;  
opoldin@hse.ru

## **An empirical analysis of students' friendship ties formation**

We consider the structure of social network of university students, and analyze factors that lead to the personal communication (friendship) network formation. Using the data from student survey and administrative information, we estimate the econometric models that assume homophily and propinquity effects on the likelihood of being friends. Specifically, we estimate linear and logistic probability models with and without fixed effects. The results confirm the significance of such friendship formation factors as sharing the same study group, living in dormitory, similar academic achievement and having the same gender. The probability of a tie increases with the number of similar attributes.

**Keywords:** social networks; friendship networks; social capital; higher education; binary choice model.

**JEL classification:** I21; Z13.

**Acknowledgements.** This research was carried out as part of the project «Institutional analysis of behavior determinants of economic agents at different markets» in the framework of Basic Research Program at the HSE in 2015. The financial support from the Government of the Russian Federation within the framework of the implementation of the 5–100 Programme Roadmap is greatly appreciated.

## **References**

Valeeva D. R., Poldin O. V., Yudkevich M. M. (2013). Friendly relationships and relationships of assistance at a university. *Educational Studies*, 4, 70–84 (in Russian).

Gavrilova A. K., Jacenko E. B. (2012). *Uchebnaja migracija iz stran SNG i Baltii: potencial i perspektivy dlja Rossii*. M.: Fond «Nasledie Evrazii» (in Russian).

Magnus J. R., Katyshev P. K., Peresetsky A. A. (2007). *Econometrics. First course*. M.: Delo (in Russian).

Radaev V. V. (2002). Term of capital, forms of capitals and their conversion. *Economic Sociology*, 3 (4), 20–32 (in Russian).

Angrist J., Pischke S. (2008). *Mostly harmless econometrics: An empiricists' companion*. Princeton: Princeton University Press.

- Baker W. (1990). Market networks and corporate behavior. *The American Journal of Sociology*, 96 (3), 589–625.
- Burt R. S. (1992). *Structural holes*. Cambridge: Harvard University Press.
- Dekker D., Krackhardt D., Snijders T. A. B. (2007). Sensitivity of MRQAP tests to collinearity and autocorrelation conditions. *Psychometrika*, 72 (4), 563–581.
- Elias J., Elias V., Ronconi L. (2007). Discrimination and social networks: Popularity among high school students in Argentina. *Inter-American Development Bank Research Network Working Paper R-539*.
- Godley J. (2008). Preference or propinquity? The relative contribution of selection and opportunity to friendship homophily in college. *Connections*, 1, 65–80.
- Granovetter M. (1973). The strength of weak ties. *American Journal of Sociology*, 78 (6), 1360–1380.
- Greene W. (2012). *Econometric analysis. 7th Edition*. Prentice Hall.
- Festinger L., Schachter S., Back K. (1950). *Social pressure in informal groups*. Stanford: Stanford University Press.
- Foster G. (2005). Making friends: A nonexperimental analysis of social pair formation. *Human Relations*, 58, 1443–1465.
- 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.
- Kenny M., Stryker S. (1996). Social network characteristics and college adjustment among racially and ethnically diverse first-year students. *Journal of College Student Development*, 37 (6), 649–658.
- Krackhardt D. (1988). Predicting with networks: Nonparametric multiple-regression analysis of dyadic data. *Social Networks*, 10 (4), 359–381.
- 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 friendship form? *The Quarterly Journal of Economics*, 121 (1), 79–119.
- McPherson M., Smith-Lovin L., Cook J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27, 415–444.
- Mayer A., Puller S. L. (2008). The old boy (and girl) network: Social network formation on university campuses. *Journal of Public Economics*, 92, 329–347.
- Preciado P., Snijders T., Burk W. J., Stattin H., Kerr M. (2012). Does proximity matter? Distance dependence of adolescent friendships. *Social Networks*, 34, 18–31.
- Soetevent A. R., Kooreman P. (2005). Social ties within school classes: The roles of gender, ethnicity, and having older siblings. *Oxford Review of Economic Policy*, 21, 373–391.
- 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.
- 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.
- Wooldridge J. (2010). *Econometric analysis of cross section and panel data. 2nd Edition*. MIT Press.

Received 16.03.2015; accepted 22.05.2015.