

Как социальные связи студентов обеспечивают эффекты обучения

Д.Р. Валеева, О. В. Польшин, М. М. Юдкевич

В работе получены оценки влияния на успеваемость отдельного студента со стороны других студентов, входящих в его микро-социальную сеть. Используются данные о третьекурсниках одного из российских университетов: ответы на вопросы о друзьях и помощниках по учебе, административные данные о результатах вступительных испытаний и успеваемости в вузе. Было установлено наличие эффектов обучения как со стороны успеваемости товарищей, так и со стороны характеристик их способностей: собственная успеваемость студента возрастала, если среди его друзей и помощников по учебе оказывались студенты с хорошей успеваемостью, высокими ЕГЭ по математике, победители или призеры олимпиад школьников.

1. Введение

Результат обучения определяется целым рядом факторов. Среди тех, которые находятся в сфере традиционного внимания исследователей, которые используются при оценке производственной функции в образовании – собственные усилия ученика, его характеристики (в том числе, личные способности и социально-демографические характеристики семьи), различные характеристики образовательной программы.

Вместе с тем, важную роль в определении результатов обучения может играть окружение учащегося – характеристики этого окружения и усилиях тех, из кого оно складывается. Такое влияние принято называть *эффектом обучения*.

В то время исследований по оценке эффектов обучения в школах достаточно много, исследования этих эффектов на уровне высшего образования пока не многочисленны и опираются на достаточно специфические данные. Так, большинство эмпирических работ по эффектам обучения в вузах опирается на анализ данных о студентах, проживающих в одной комнате или блоке в студенческом общежитии (см., например, Sacerdote, 2001, Zimmerman, 2003, Brunello et al., 2010). Были найдены значимые эффекты обучения, причем в двух первых работах они оказались нелинейными – влияние окружения проявлялось неодинаково как на студентов, так и со стороны студентов с различными способностями.

Эффекты обучения собственно со стороны студенческой группы в силу эндогенного характера формирования индивидуальных учебных программ в большинстве вузов практически не анализировались. Исключение составляют работы, выполненные на данных военных академий США, где взаимодействие среди студентов внутри отдельных подразделений весьма интенсивно, а группы и учебные программы формируются административным образом, не оставляя учащимся свободы выбора курсов (Lyle, 2007, 2009, Carrell, Fullerton, West 2009). В этих работах исследовалось влияние на успеваемость студента со стороны характеристик студенческой группы, сформированной случайным образом на 1 курсе.

В работе Андрушака, Польшина и Юдкевич (2012) установлено наличие эффектов обучения в административно формируемой студенческой группе. Оценки по отдельным дисциплинам и агрегированная успеваемость на первом курсе обучения у отдельного студента возрастали с ростом способностей одноклассников, причем подобное влияние носило нелинейный характер: чем выше относительное число способных студентов в группе, тем лучше оказывалась успеваемость студента, также относящегося к числу наиболее подготовленных.

Изучение влияния со стороны группы в целом имеет определенные недостатки. Так, считается, что всё члены группы имеют одинаковое влияние на студента, что является сильно упрощающим предположением. Также при анализе групповых эффектов трудно разграничить влияние характеристик окружения и влияние поведение окружения. Это разграничение очень важно как для понимания механизма действия эффектов обучения, так и для образовательной политики.

Многие параметры окружения (способности, социально-демографические характеристики) фиксированы и не подвержены изменению. Поведение (успеваемость) окружения может изменяться под воздействием образовательной политики. Стимулирование достижений части студентов может оказывать благоприятное влияние на успехи окружения этих студентов. Возросшая успеваемость окружения, в свою очередь, способна усилить первоначальный эффект политики. Так возникает положительная обратная связь, эффект социального мультипликатора, при котором прямое изменение успеваемости от внешнего воздействия на студента дополняется косвенным эффектом от его социальных связей.

Преодолеть ограничения, накладываемые анализом групповых взаимодействий, позволяет анализ индивидуальных связей каждого студента. В ряде недавних работ рассматривался вопрос о том, какое влияние социальные связи оказывают на поведение и успеваемость учащихся.

Мейер и Пуллер (Mayer и Puller (2008)) опирались на данные о студентах одного американского университета, полученные из социальной сети Facebook. Среди прочих результатов, было обнаружено, что успеваемость студента сильно коррелирует со средней успеваемостью друзей: при увеличении среднего балла друзей на единицу собственный средний балл возрастает на 0.46.

Лави и Сэнд (Lavy, Sand 2012) рассматривали влияние различных по тесноте дружеских отношений на учебу и поведение израильских школьников. Наибольшее положительное влияние на успеваемость со стороны ровесников возникает в ситуации, характеризующейся взаимными дружескими связями.

Эмпирические исследования в работах Calvo-Armengol и др. (2009), Bramoullé et al. (2009) и Lin (2010) базируются на анализе обширной выборочной базы данных о здоровье школьников старших классов в США (The National Longitudinal Study of Adolescent Health – Add Health), содержащей среди прочего сведения о дружеских отношениях среди респондентов. В исследовании Calvo-Armengol и др. (2009) показано, что число друзей является важным положительным фактором успеваемости. Брамулле и др. (Bramoullé и др. 2009) обнаружили значимо положительное влияние друзей на уровень и виды активности школьника в свободное от учебы время. В работе Lin (2010) найдено, что и сама успеваемость друзей, и их социально-экономические характеристики значимо связаны с успеваемостью учащегося. В двух последних работах использован аппарат пространственного регрессионного анализа, позволивший точнее оценить механизм действия эффектов сообучения.

Во всех обсуждающихся выше работах по изучению эффектов сообучения в социальных сетях использовались подробные сведения о структуре социальных связей учащихся применительно к среднему образованию. В высшем образовании исследования базировались на сведениях из интернета или были основаны на данных об обмене электронными письмами, что позволяет лишь приближенно оценить истинные дружеские связи. В настоящей работе используются данные о социальных связях в вузе, полученные путем прямого анкетирования студентов третьего курса одного российского университета. Помимо дружеских связей, из анкет были получены данные о том, к кому студенты обращаются за помощью по учебным вопросам.

Нами обнаружены значимые положительные эффекты сообучения как со стороны успеваемости, так и со стороны способностей товарищей. В то же время влияние одноклассников в целом оказалось не значимым. Полученные в настоящей работе результаты позволяют лучше понять механизмы действия эффекта сообучения в группе, ранее обнаруженные в зарубежных и российских исследованиях.

2. Эмпирические модели эффектов сообучения

В эмпирических исследованиях в качестве основной модели рассматривается следующая её спецификация:

$$y_i = \alpha + \beta \bar{y}_{-i}^{peer} + \gamma' \mathbf{x}_i + \delta' \bar{\mathbf{x}}_{-i}^{peer} + \varepsilon_i, i = 1, 2, \dots, n., \quad (1)$$

где y_i – показатель успеваемости студента i , \mathbf{x}_i – $k \times 1$ вектор индивидуальных характеристик студента i , $\bar{\mathbf{x}}_{-i}^{peer}$ – $k \times 1$ вектор средних экзогенных характеристик группы студентов, референтной для студента i , формирующей его окружение, \bar{y}_{-i}^{peer} – средняя успеваемость референтной группы студента i , ε_i – случайные возмущения, n – число студентов в выборке.

Коэффициент β в модели (1) измеряет, как индивидуальная успеваемость зависит от успеваемости окружения (эндогенный эффект). Компоненты вектора δ измеряют влияние экзогенных характеристик окружения (экзогенный, или контекстуальный, эффект)¹. Важное отличие между этими воздействиями состоит в том, что только эндогенный эффект ($\beta \neq 0$) способен мультиплицироваться: если в результате некоторого внешнего воздействия (ε_i в уравнении (1)) успеваемость одного студента улучшается, то это влияет на его окружение. Изменение достижений окружения, в свою очередь, вызывает ответное воздействие на студента, и т.д. Чисто экзогенные эффекты сообучения ($\delta \neq 0, \beta = 0$) не обладают данным свойством, так как внешние шоки не влияют на фиксированные характеристики студентов.

Присутствие эндогенных эффектов имеет большое значение для образовательной политики. Например, дополнительные занятия с частью студентов могут привести к лучшей успеваемости не только тех студентов, кто посещает эти дополнительные занятия, но вызовут рост достижений их товарищей.

Из-за того, что не только успеваемость окружения студента влияет на его академические показатели, но и его собственные успехи влияют на товарищей, оценивание коэффициентов модели (1) методом наименьших квадратов дает смещенные оценки. С целью избежать смещения оценок часто используется приведенная форма модели:

$$y_i = \alpha + \gamma' \mathbf{x}_i + \delta' \bar{\mathbf{x}}_{-i}^{peer} + \varepsilon_i. \quad (2)$$

Она дает возможность оценить общий эффект сообучения, без различения эндогенных и экзогенных эффектов.

Наличие данных о социальных связях позволяет оценить отдельно эндогенные и экзогенные эффекты сообучения. Используемая для этого модель аналогична авторегрессионной модели пространственного взаимодействия, в которой учитывается географическая близость между различными объектами в пространстве.

Логично предположить, что именно влияние близких товарищей оказывает наибольшее влияние на успеваемость студента, причем каждый студент может иметь свой специфический круг общения. Тогда средние величины из (1) можно записать в виде:

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

где усреднение проведено по множеству P_i , в которую входят n_i товарищей студента i . При расчете средних величин показатели самого студента не учитываются.

В матричной форме модель (1) для выборки из n студентов с учетом (3) записывается в виде:

$$\mathbf{y} = \alpha \mathbf{i} + \beta \mathbf{G} \mathbf{y} + \mathbf{X} \gamma + \mathbf{G} \mathbf{X} \delta + \boldsymbol{\varepsilon}, \quad E[\boldsymbol{\varepsilon} | \mathbf{X}] = 0, \quad (4)$$

где – \mathbf{y} – $n \times 1$ вектор успеваемости n студентов, \mathbf{i} – $n \times 1$ единичный вектор, \mathbf{X} – $n \times k$ матрица характеристик студентов, \mathbf{G} – $n \times n$ матрица взаимодействий с элементами

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

¹ Мы пренебрегаем влиянием ненаблюдаемых характеристик, общих для студента и его окружения.

Диагональные элементы \mathbf{G} равны нулю, она нормирована по строкам. Строка i матрицы \mathbf{GX} состоит из средних значений k характеристик товарищей студента i , а i -й элемент вектора \mathbf{Gy} есть средняя успеваемость товарищей студента i :

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

Модель (4) по структуре похожа авторегрессионную модель пространственного взаимодействия (LeSage, Pace (2009) классифицируют данную модель как Spatial Durbin Model), которая может быть оценена методом максимального правдоподобия.

В пространственной статистике аналогом матрицы \mathbf{G} является матрица пространственных весов, которая описывает географическую близость между различными объектами в пространстве и поэтому является симметричной. В случае социальных связей матрица \mathbf{G} может не быть симметричной, если студент i считает студента j своим другом, но не наоборот.

Проблема оценивания вектора параметров $\alpha, \beta, \gamma, \delta$ в (4) связана с эндогенностью слагаемого \mathbf{GY} в правой части. Как показали Bramoullé, Djebbari и Fortin (2009), параметры идентифицируемы при условии, что матрицы \mathbf{I}, \mathbf{G} , and \mathbf{G}^2 линейно независимы. При сетевых взаимодействиях линейная независимость обеспечивается присутствием в сетевой структуре нетранзитивных триад², Это требование обычно выполняется при рассмотрении социальных связей, так как обычно не все друзья друзей являются друзьями студента.

Как было сказано выше, изменение объясняющей переменной для одного студента влияет как на его собственную успеваемость (прямое воздействие), так и опосредованно на успеваемости других студентов (косвенное воздействие). Оценки коэффициентов пространственной (сетевой) регрессии учитывают взаимозависимость между наблюдениями, которые представляют отдельных студентов. Это свойство выгодно отличает данную модель от обычной регрессионной модели, в которой наблюдения независимы друг от друга.

В Приложении описаны некоторые количественные показатели, характеризующие количественную связь зависимой и независимой переменной, отражающие специфику сетевых взаимодействий. Среднее прямое влияние характеризует воздействие на студента со стороны его собственных характеристик, усредненное по выборке. Среднее полное влияние на студента интерпретируется как воздействие на успеваемость одного студента со стороны характеристик всех студентов. Среднее полное влияние от студента имеет смысл влияния характеристики одного студента на всех студентов. Среднее косвенное влияние, по определению, есть разница между средним полным и средним прямым влиянием. Ниже будут приведены расчеты данных показателей для рассматриваемой сети.

3. Оценивание эффектов обучения в социальной сети

Данные

В исследовании были использованы данные об успеваемости и характеристиках студентов факультета экономики НИУ ВШЭ, обучавшихся в 2011-2012 учебном году на 3 курсе бакалавриата. В качестве показателя успеваемости использовался средний балл за первые три года обучения. В Высшей школе экономики используется 10-бальная шкала оценок — чем выше балл, тем выше достижения по предмету.

² Нетранзитивной триадой называют ситуацию, при которой студент А связан со студентом Б, студент Б связан со студентом В, но студенты Б и В непосредственно не связаны друг с другом.

Экзогенными характеристиками, влияющими на успеваемость студента, выступили баллы ЕГЭ по математике и русскому языку, принадлежность к победителям и призерам Всероссийской олимпиады школьников или Межрегиональной олимпиады ВШЭ³.

Описательные статистики переменных приведены в таблице 1.

Таблица 1. Описательные статистики

Переменная	Число наблюдений	Среднее значение	Стандартное отклонение	Минимальное значение	Максимальное значение
<i>Друзья/ Помощники по учебе</i>					
ЕГЭ по русскому языку	167/171	80,0	9,6	50	100
ЕГЭ по математике	167/171	76,4	8,0	48	100
Победитель/призер Всероссийской олимпиады	167/171	0,16	0,37	0	1
Победитель/призер Межрегиональной олимпиады	167/171	0,24	0,43	0	1
Средний балл за 3 курса	167/171	7,35	0,88	5,57	9,21

Для определения референтной группы был проведен опрос, с просьбой указать в анкете до пяти однокурсников, с которыми студент общается больше всего («друзья»), и до пяти однокурсников, к которым студент чаще всего обращается за помощью в учебных вопросах («помощники по учебе»). После обработки анкет были сформированы матрицы социальных связей каждого студента двух типов – для друзей и помощников по учебе. Доля друзей среди помощников по учебе составляет 60%.

На рис. 1 приведена диаграмма рассеяния, связывающая собственный средний балл студента и усредненный средний балл его друзей. Просматривается четко выраженная корреляция (коэффициент корреляции равен 0,42).

³ Эти показатели являются значимыми предикторами академических успехов студентов (Польдин 2011, Давтян, Пересецкий 2011).

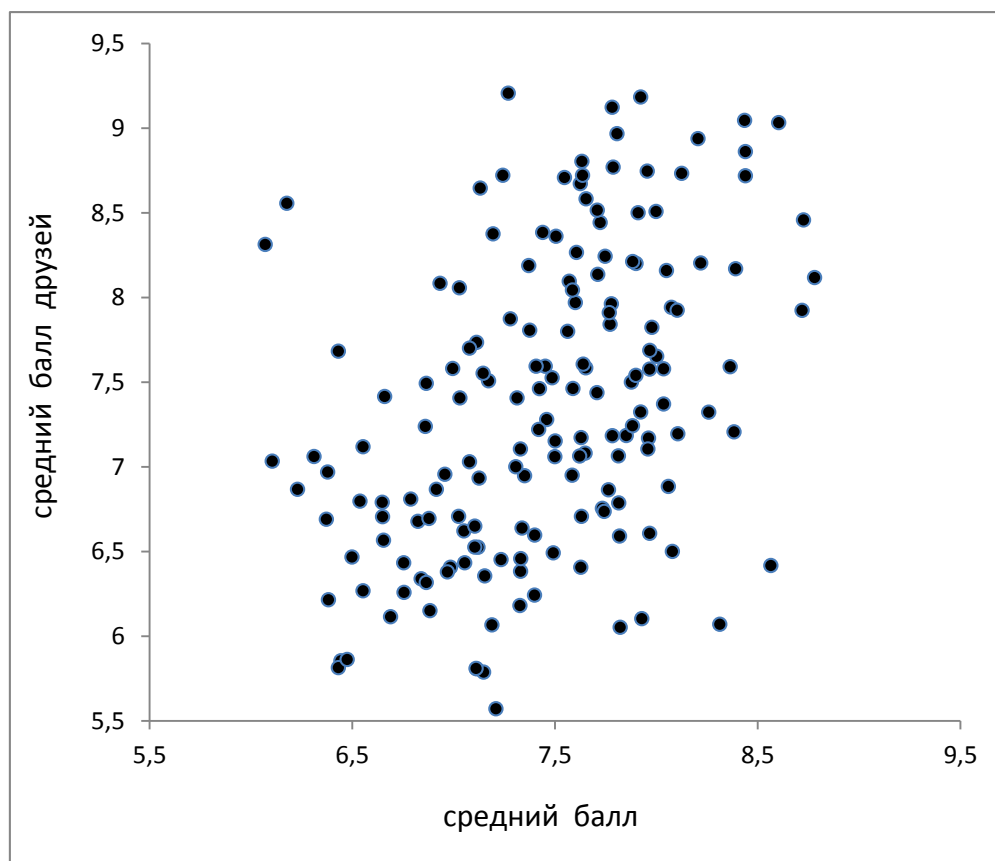


Рис 1. Диаграмма рассеяния – средний балл студента и средний балл его друзей.

В Таблице 2 приведены некоторые характеристики сети. Все значения измеряются в пределах от 0 до 1. Сети помощи более централизованы, чем сети дружбы, что можно объяснить тем, что за помощью обращаются чаще к одним и тем же студентам. Сеть дружбы в большей степени взаимна, чем сети помощи — почти половина всех возможных связей взаимна.

Таблица 2. Описательная статистика по сетям.

	Сеть помощи	Сеть дружбы
Плотность	0.02	0.02
Централизация (входящая)	0.10	0.06
Реципрокность	0.19	0.42
Коэффициент кластеризации	0.24	0.27

Примечание: Плотность сети есть наблюдаемая доля связей (как входящих, так и исходящих) в сети по отношению к общему возможному количеству связей в этой сети. Централизация (входящая) – наблюдаемая доля исходящих связей в сети по отношению к общему возможному количеству исходящих связей в этой сети. Реципрокность – наблюдаемая доля взаимных связей в сети по отношению к ситуации, когда все связи в этой сети были бы взаимными. Коэффициент кластеризации – коэффициент, показывающий вероятность образования связи между участниками сети, которые связаны с одним общим участником.

Результаты оценивания

Нами оценивались как модель в приведенной форме, позволяющая обнаружить влияние окружения, так и модель в полной спецификации, дающая возможность различить эндогенные

и контекстуальные эффекты сообучения. Помимо полного набора экзогенных параметров, характеризующих способности (баллы ЕГЭ по математике и русскому языку, статус победителя и призера олимпиад), также рассматривались спецификации, в которых использовались только баллы ЕГЭ.

В таблице 3 приведены оценки модели методом наименьших квадратов для друзей в приведенной форме:

$$y = \alpha_r i + Xy_r + GX\delta_r + \varepsilon_r. \quad (6)$$

Собственные характеристики студента значимы на 1% уровне. Влияние ЕГЭ по русскому языку слабее влияния ЕГЭ по математике: прирост ЕГЭ по математике на 10 баллов приводит к увеличению среднего балла в вузе на 0,41, по русскому языку – на 0,23. Статус победителя/призера Всероссийской Олимпиады дает ощутимо большую прибавку (0,81 балла) к успеваемости, чем статус победителя/призера Межрегиональной Олимпиады ВШЭ (0,43 балла).

В этой модели коэффициент δ_r «смешивает» в себе эндогенные и экзогенные эффекты, не позволяя идентифицировать их по отдельности. Значимыми оказались коэффициенты при переменных «средний балл ЕГЭ друзей по математике» и «среднее число среди друзей победителей/призеров Всероссийской Олимпиады». Таким образом, можно говорить о присутствии эффектов сообучения: собственная успеваемость студента возрастает при увеличении ЕГЭ по математике у его друзей и в случае, когда среди них есть победители или призеры олимпиад школьников.

Также были построены симметричные матрицы связей: два студента в ней считались друзьями, если номинация была взаимной. Для симметричных связей коэффициент при ЕГЭ друзей по математике значим уже на 1% уровне.

Оценки модели в полной спецификации с помощью пространственной авторегрессионной модели (4) методом максимального правдоподобия приведены в таблице 4а. Оценки коэффициентов β значимо положительные. Прямой эффект от прироста среднего балла у друзей на единицу состоит в росте собственного среднего балла студента на 0,25. Коэффициенты при экзогенных переменных оказались незначимыми. Данный результат может свидетельствовать о том, что на успехи студента более влияет именно фактическая успеваемость его друзей, а не их экзогенные характеристики.

В таблице 4б приведены оценки среднего прямого, среднего косвенного и среднего полного влияния экзогенных характеристик для сети друзей для случая несимметричных связей.

Студенты, связанные дружескими связями, обычно учатся в одной и той же студенческой группе формируемой административно для проведения семинарских занятий. Для сравнения эффекта сообучения от друзей и эффекта от всей группы, было проведено оценивание моделей (1) и (2) для ситуации, при которой в качестве референтной группы выступали все студенты из студенческой группы, принявшие участие в анкетном опросе. Результаты показаны в таблице 5. Влияние экзогенных характеристик студенческой группы оказалось положительно, но слабее по величине по сравнению с друзьями и статистически незначимо. Эндогенный эффект отсутствует.

В данной работе нас интересует влияние окружения на академические достижения студента. Именно поэтому при анкетировании помимо друзей предлагалось также назвать однокурсников, к которым студент обращается за помощью в учебных вопросах. Списки друзей и помощников по учебе частично пересекаются, но в целом студенты по понятной причине предпочитают консультироваться по учебным вопросам с однокурсниками, имеющими успеваемость выше собственной. Как и в случае с друзьями, существует положительная корреляция между средним баллом студента и средним баллом его помощников по учебе.

В таблице 6а приведены результаты оценивания для случая сетей из помощников по учебе. Качественно результаты похожи на оценки из таблиц 3 и 4 для дружеских сетей. Коэффициент детерминации для сети помощи чуть выше, чем для сети дружбы. Оценки показывают присутствие статистически значимых эффектов сообучения. Чем выше

фиксированные характеристики способностей и достижения консультантов студента по учебным вопросам, тем лучше успеваемость и самого студента. Величина эндогенного эффекта для сети помощи практически совпадает с результатом для сети дружбы. Экзогенные эффекты от ЕГЭ помощников по математике, присутствия среди них победителей/призеров Всероссийской Олимпиады и Межрегиональной Олимпиады ВШЭ оказались также статистически значимы.

В таблице 6б приведены оценки среднего прямого, среднего косвенного и среднего полного влияния экзогенных характеристик для сети помощников.

Для тех же студентов на первом году обучения, были найдены значимые эффекты сообучения в приведенной форме модели, проявляющиеся в положительном влиянии на общую успеваемость сильных студентов со стороны хорошо подготовленных студентов в административно сформированных группах (Андрущак, Польшин и Юдкевич, 2012). На втором и третьем курсах влияние группы становится незначимым. Полученные в настоящей работе результаты позволяют сделать некоторые предположения относительно механизма действия подобного эффекта.

Можно предположить, что сильные студенты в группе выступают примером для подражания для других членов группы и способны консультировать одноклассников по учебным вопросам. Это объясняет нелинейный характер проявления эффектов сообучения на первом курсе. С течением времени возникают более узкие дружеские связи между студентами с близким уровнем подготовки. Друзья оказывают влияние на отношение к учебе, от них легче получить помощь по учебным вопросам. Влияние остальных членов группы ослабевает, что объясняет отсутствие значимых эффектов сообучения от группы на втором и третьем курсах. При этом влияние товарищей остается значимым.

В таблице 7 сведены воедино качественные результаты оценивания.

Таблица 7. Присутствие эффектов сообучения в моделях с различной спецификацией референтной группы.

	Референтная группа		
	Студенческая группа	Друзья	Помощники по учебе
Эффект сообучения (без различения эндогенных и экзогенных эффектов)	нет	да	да
Эндогенный эффект	нет	да	да
Экзогенный эффект	нет	нет	да

4. Заключение

Изучение влияния социального взаимодействия учащихся на их академические достижения представляет значительный интерес, как для понимания природы образовательного процесса, так и с точки зрения повышения результативности обучения. Оценивание эффектов сообучения на практике связано с рядом трудностей. Одна из проблем состоит в корректной идентификации группы учащихся, взаимодействующих при обучении и способных оказывать влияние друг на друга. Важно также разграничить влияние текущих достижений студентов, поддающихся изменениям, и их неизменных характеристик, обусловленных способностями, социальным статусом и прочими обстоятельствами.

В настоящей работе использованы данные о социальных связях среди студентов третьего курса факультета экономики одного из ведущих российских вузов. Путем опроса были получены данные о том, с кем из одноклассников студенты поддерживают дружеские отношения и к кому они обращаются за помощью по учебным вопросам. В результате для

каждого студента был сформирован список учащихся, имеющих оказывающих наибольшее влияние в процессе обучения.

Информация об индивидуальных социальных связях каждого студента позволило применить для анализа данных пространственную регрессионную модель и с её помощью решить задачу о раздельном оценивании влияния текущей успеваемости товарищей (эндогенного эффекта сообучения) и характеристик их способностей (экзогенного эффекта).

Нами обнаружены значимые положительные эффекты сообучения как со стороны успеваемости, так и со стороны способностей студентов. При увеличении среднего балла у товарищей по учебе на единицу происходит прирост среднего балла у самого студента на одну четверть единицы.

Присутствие эндогенных эффектов имеет значение для образовательной политики. В этом случае стимулирование успеваемости части студентов имеет мультипликационный эффект: рост успехов отдельных студентов вызывает улучшение успеваемости их товарищей, что, в свою очередь благоприятно влияет на успехи тех, с кем они дружат и взаимодействуют. Использование взаимного обучения в студенческих группах через дискуссии, совместное решение задач и проектной работе позволит существенно повысить эффективность освоения знаний.

Приложение. Особенности интерпретации коэффициентов пространственной регрессионной модели

Изменение объясняющей переменной для одного студента влияет как на его собственную успеваемость (прямое воздействие), так и опосредованно на успеваемости других студентов (косвенное воздействие). Оценки коэффициентов пространственной (сетевой) регрессии учитывают взаимозависимость между наблюдениями, которые представляют отдельных студентов.

Количественная связь зависимой и независимой переменной описывается величиной маржинального (предельного) эффекта. Следуя (LeSage, Pace 2009), можно получить выражения для маржинальных эффектов в сетевой регрессии.

Преобразуя выражение (4), получаем

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

где

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

\mathbf{x}_j – $n \times 1$ вектор равный j -му столбцу матрицы \mathbf{X} , его элементами являются значения j -й экзогенной переменной для n студентов.

Для успеваемости i -го студента из (П1) следует:

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

Влияние j -й экзогенной переменной на y_i состоит из суммы

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

В этом выражении отчетливо прослеживается влияние, которому подвергается отдельный студент со стороны характеристик других студентов. Маржинальный эффект от изменения непрерывной переменной измеряется частной производной

$$\frac{\partial y_i}{\partial x_j} = (S_j)_{ii}. \quad (ПЗ)$$

Для дискретной независимой переменной маржинальный эффект определяется через конечную разность:

$$\frac{\Delta y_i}{\Delta x_j} = [y_i | x_j = 1] - [y_i | x_j = 0].$$

Диагональные элементы матрицы S_j описывают прямое воздействие на успеваемость студента со стороны его собственных характеристик. Они не равны параметрам γ_j (см. выражение (П2)), так как учитывают обратную связь – способности студента A влияют на его успеваемость, она влияет (при $\beta \neq 0$) на успеваемость других студентов, связанных сетью с A , а их успеваемость воздействует обратно на A .

Недиагональные элементы представляют собой косвенное воздействие характеристик одного студента (j) на успеваемость другого студента (l). Это влияние существует не только между друзьями, оно распространяется на других однокурсников – друзей друзей и т.д.

Элементы S_j различаются между собой, что отражает различие положений студентов в структуре сетевых связей. LeSage и Pace (2009) предложили скалярные агрегированные показатели, измеряющие усредненное влияние экзогенных переменных.

Среднее прямое влияние рассчитывается как усредненное значение элементов главной диагонали матрицы S_j :

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

Среднее полное влияние на наблюдение (студента) есть среднее значение суммы рядов матрицы S_j , интерпретируется как изменение успеваемости одного студента при увеличении j -й независимой переменной на единицу у всех студентов. Среднее полное влияние от наблюдения (студента) есть среднее значение суммы столбцов матрицы S_j , имеет смысл приращения успеваемости у всех студентов при увеличении j -й независимой переменной на единицу у одного студента. Хотя разные по смыслу, эти две характеристики количественно равны друг другу:

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

Среднее косвенное влияние есть, по определению, разница между средним полным и средним прямым влиянием:

$$AIM_j = ATI_j - ADI_j.$$

Литература

- Андрущак Г.В., Польдин О.В., Юдкевич М.М. (2012). Эффекты сообучения в административно формируемых студенческих группах // Прикладная эконометрика, 2012. № 2. С. 3—16
- Пересецкий А. А., Давтян М. А. (2011). Эффективность ЕГЭ и олимпиад как инструмента отбора абитуриентов. // Прикладная эконометрика, 23 (3), 41–56.
- Польдин О. В., Юдкевич М. М. (2011) Эффекты сообучения в высшем образовании: обзор теоретических и эмпирических подходов // Вопросы образования, 4, 106–123.

- Польдин О. В. (2011). Прогнозирование успеваемости в вузе по результатам ЕГЭ. // Прикладная эконометрика, 21 (1), 56–69.
- Borgatti S.P., Everett M.G., Freeman L.C. (2002). Ucinet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies
- Bramoullé Y., Djebbari H., Fortin B. (2009). Identification of peer effects through social networks. *Journal of Econometrics* 150, 41–55.
- Brunello G., De Paola M., Scoppa V. (2010). Peer effects in higher education: Does the field of study matter? *Economic Inquiry*, 48 (3), 621–634.
- Calvó-Armengol A, Patacchini E, Zenou Y. (2009). Peer effects and social networks in education. *Review of Economic Studies*, 76, 1239-1267/
- Carrell S., Fullerton R., West J. (2009). Does your cohort matter? Measuring peer effects in college achievement. *Journal of Labor Economics*, 27 (3), 439–464.
- De Giorgi G., Pellizzari M., Redaelli S. (2010) Identification of Social Interactions through Partially Overlapping Peer Groups. *American Economic Journal: Applied Economics*, 2 (2). 241-275
- Laumann E.O., Guttman L. (1966). The relative associational contiguity of occupations in an urban setting. *American Sociological Review* 31, 169–178.
- Lavy V., Sand E. (2012). The Friends Factor: How students' social networks affect their academic achievement and well-being. Working paper.
- Lee L.F. (2003). Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances. *Econometric Reviews*. 22 (4), 307–335.
- LeSage J.P., Pace R.K. (2009). Introduction to spatial econometrics. Boca Raton, US: CRC Press Taylor & Francis Group.
- Lin Xu (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., Puller S. (2008). The old boy (and girl) network: Social network formation on university campuses. *Journal of Public Economics*, 92, 329–347
- R Development Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>
- Sacerdote B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics*, 116, 681–704.
- Wasserman S., Faust K. (1994). Social Network Analysis: Methods and Applications. Cambridge, ENG and New York: Cambridge University Press.
- Zimmerman D. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and Statistics*, 85 (1), 9–23.

Таблица 3. Оценки эффектов сообучения для друзей (без различения эндогенных и экзогенных эффектов).

Переменная	Несимметричные связи		Взаимные связи	
	(1)	(2)	(3)	(4)
Константа	0,075 (0,063)	0,445 (0,338)	0,239 (0,258)	0,404 (0,384)
ЕГЭ студента по русскому языку	0,023*** (4,063)	0,021*** (3,445)	0,024*** (4,306)	0,022*** (3,437)
ЕГЭ студента по математике	0,041*** (6,097)	0,047*** (6,331)	0,036*** (5,180)	0,044*** (5,789)
Статус победителя/призера Всероссийской Олимпиады	0,808*** (5,361)		0,843*** (5,399)	
Статус победителя/призера Межрегиональной Олимпиады ВШЭ	0,429*** (3,465)		0,479*** (3,838)	
Средний ЕГЭ друзей по русскому языку	0,002 (0,181)	-0,011 (-0,905)	-0,004 (-0,491)	-0,008 (-0,959)
Средний ЕГЭ друзей по математике	0,024* (1,817)	0,032** (2,346)	0,032*** (3,222)	0,032*** (3,062)
Среднее число среди друзей победителей/призеров Всероссийской Олимпиады	0,545** (2,213)		0,302 (1,417)	
Среднее число среди друзей победителей/призеров Межрегиональной Олимпиады ВШЭ	0,303 (1,371)		-0,073 (-0,400)	
Число наблюдений	167	167	156	156
R ²	0,47	0,32	0,49	0,32

*, **, *** – значимость на 10%, 5% и 1% уровне, соответственно. В скобках указаны t-статистики.

Таблица 4а. Оценки эффектов сообучения для друзей (с различием эндогенных и экзогенных эффектов).

Переменная	Несимметричные связи		Взаимные связи	
	(1)	(2)	(3)	(4)
Константа	-0,066 (-0,059)	0,140 (0,115)	-0,215 (-0,249)	-0,119 (-0,123)
ЕГЭ студента по русскому языку	0,024*** (4,497)	0,022*** (3,865)	0,025*** (4,812)	0,023*** (4,000)
ЕГЭ студента по математике	0,038*** (5,971)	0,043*** (6,219)	0,032*** (5,004)	0,040*** (5,621)
Статус победителя/призера Всероссийской Олимпиады	0,772*** (5,443)		0,806*** (5,563)	
Статус победителя/призера Межрегиональной Олимпиады ВШЭ	0,416*** (3,572)		0,448*** (3,876)	
Средний ЕГЭ друзей по русскому языку	-0,005 (-0,470)	-0,015 (-1,428)	-0,009 (-1,205)	-0,011 (-1,462)
Средний ЕГЭ друзей математике	0,011 (0,850)	0,012 (0,864)	0,024** (2,492)	0,018* (1,786)
Среднее число среди друзей победителей/призеров Всероссийской Олимпиады	0,237 (0,961)		0,075 (0,363)	
Среднее число среди друзей победителей/призеров Межрегиональной Олимпиады ВШЭ	0,171 (0,802)		-0,203 (-1,183)	
β (эндогенный эффект)	0,253*** (2,612)	0,339*** (3,662)	0,232*** (3,172)	0,277*** (3,860)
Число наблюдений	167	167	156	156
R ²	0,48	0,32	0,48	0,32

*** – значимость на 1% уровне. В скобках указаны t-статистики.

Таблица 4б. Среднее прямое, косвенное и полное влияние экзогенных характеристик для сети друзей (для моделей (1) и (2) в таблице 4а).

	(1)			(2)		
	Прямое влияние	Косвенное влияние	Полное влияние	Прямое влияние	Косвенное влияние	Полное влияние
ЕГЭ студента по русскому языку	0,024***	0,002	0,025*	0,021***	-0,011	0,010
ЕГЭ студента по математике	0,039***	0,027	0,065***	0,044***	0,039**	0,083***
Статус победителя/призера Всероссийской Олимпиады	0,793***	0,560*	1,353***			
Статус победителя/призера Межрегиональной Олимпиады ВШЭ	0,428***	0,369	0,797**			

*, **, *** – значимость на 10%, 5% и 1% уровне, соответственно

Таблица 5. Оценки эффектов сообучения для одноклассников.

Переменная	Без эндогенной переменной		С эндогенной переменной	
	(1)	(2)	(3)	(4)
Константа	-0,842 (-0,233)	3,164 (0,946)	-0,895 (-0,254)	2,928 (0,870)
ЕГЭ студента по русскому языку	0,025*** (4,243)	0,020*** (3,235)	0,025*** (4,377)	0,021*** (3,295)
ЕГЭ студента по математике	0,040*** (5,720)	0,048*** (6,281)	0,040*** (5,892)	0,048*** (6,377)
Статус победителя/призера Всероссийской Олимпиады	0,906*** (6,057)		0,912*** (6,244)	
Статус победителя/призера Межрегиональной Олимпиады ВШЭ	0,478*** (3,680)		0,480*** (3,799)	
Средний ЕГЭ друзей по русскому языку	0,024 (0,468)	-0,021 (-0,516)	0,029 (0,558)	-0,021 (-0,518)
Средний ЕГЭ друзей по математике	0,008 (0,258)	0,008 (0,297)	0,015 (0,430)	0,004 (0,134)
Среднее число среди друзей победителей/призеров Всероссийской Олимпиады	1,017 (1,250)		1,220 (1,306)	
Среднее число среди друзей победителей/призеров Межрегиональной Олимпиады ВШЭ	0,520 (0,575)		0,616 (0,673)	
β (эндогенный эффект)			-0,119 (-0,441)	-0,073 (-0,322)
Число наблюдений	167	167	167	167
R ²	0,44	0,29	0,44	0,29

*** – значимость на 1% уровне. В скобках указаны t-статистики.

Таблица 6а. Оценки эффектов сообучения для помощников по учебе.

Переменная	Без эндогенной переменной		С эндогенной переменной	
	(1)	(2)	(3)	(4)
Константа	-1,904 (-1,650)	-1,059 (-0,836)	-2,107* (-1,933)	-1,554 (-1,321)
ЕГЭ студента по русскому языку	0,024*** (4,547)	0,022*** (3,605)	0,024*** (4,756)	0,021*** (3,842)
ЕГЭ студента по математике	0,034*** (5,210)	0,039*** (5,419)	0,032*** (5,160)	0,037*** (5,387)
Статус победителя/призера Всероссийской Олимпиады	0,857*** (6,100)		0,821*** (6,188)	
Статус победителя/призера Межрегиональной Олимпиады ВШЭ	0,440*** (3,699)		0,427*** (3,801)	
Средний ЕГЭ помощников по русскому языку	0,011 (1,026)	-0,004 (-0,302)	0,005 (0,464)	-0,010 (-0,919)
Средний ЕГЭ помощников по математике	0,043*** (3,969)	0,051*** (4,221)	0,031*** (2,668)	0,033*** (2,564)
Среднее число среди помощников победителей/призеров Всероссийской Олимпиады	0,625** (2,482)		0,380 (1,517)	
Среднее число среди помощников победителей/призеров Межрегиональной Олимпиады ВШЭ	0,407** (1,983)		0,235 (1,156)	
β (эндогенный эффект)			0,250** (2,481)	0,337*** (3,520)
Число наблюдений	171	171	171	171
R ²	0.52	0.35	0,51	0,31

, * – значимость на 5% и 1% уровне, соответственно. В скобках указаны t-статистики.

Таблица 6б. Среднее прямое, косвенное и полное влияние экзогенных характеристик для сети помощников (для моделей (3) и (4) в таблице 6а).

	(3)			(4)		
	Прямое влияние	Косвенное влияние	Полное влияние	Прямое влияние	Косвенное влияние	Полное влияние
ЕГЭ студента по русскому языку	0,026***	0,025	0,050	0,021***	-0,005	0,017
ЕГЭ студента по математике	0,040***	0,008	0,047	0,038***	0,067***	0,106***
Статус победителя/призера Всероссийской Олимпиады	0,907***	0,947	1,854			
Статус победителя/призера Межрегиональной Олимпиады ВШЭ	0,477***	0,514	0,991**			

*, **, *** – значимость на 10%, 5% и 1% уровне, соответственно