

---

---

## **ПРАКТИКИ СБОРА И АНАЛИЗА ФОРМАЛИЗОВАННЫХ ДАННЫХ**

Т.Е. Хавенсон, Ю.А. Тюменева  
(Москва)

### **ПЕРЕХОД ОТ КОРРЕЛЯЦИОННОГО К КВАЗИЭКСПЕРИМЕНТАЛЬНОМУ ПЛАНУ В ИССЛЕДОВАНИЯХ ШКОЛЬНОГО ОБРАЗОВАНИЯ: ОПЫТ ПРИМЕНЕНИЯ РЕГРЕССИЙ С ФИКСИРОВАННЫМИ ЭФФЕКТАМИ<sup>1</sup>**

В статье предпринята попытка выявить причинно-следственные связи характеристик учителей и успеваемости школьников на основе срезовых данных международного исследования школьного образования TIMSS (волна 2007). Подробно описана методология, основанная на регрессионных моделях с фиксированными эффектами, а именно на оценках первой разницы, позволяющая перейти к квазиэкспериментальному плану исследования. Сравниваются результаты, полученные с использованием линейного регрессионного анализа, с результатами, полученными путем применения данной методологии.

*Ключевые слова:* квазиэкспериментальный план исследования, срезовые данные, регрессионный анализ, регрессионные модели с фиксированными эффектами, школьная успеваемость.

---

**Татьяна Евгеньевна Хавенсон** – старший преподаватель кафедры методов сбора и анализа социологической информации факультета социологии НИУ ВШЭ; младший научный сотрудник Международной лаборатории анализа образовательной политики НИУ ВШЭ. E-mail: tkhavenson@hse.ru.

**Юлия Алексеевна Тюменева** – кандидат психологических наук, старший научный сотрудник Международной лаборатории анализа образовательной политики НИУ ВШЭ. E-mail: jutu@yandex.ru.

## *Введение*

В социальных науках мы часто имеем дело с данными, собранными в рамках срезовых исследований, т. е. в один момент времени на одной выборке респондентов. Такие исследования преобладают по сравнению с лонгитюдными или экспериментальными. Так, проведенный Т. Зиделем и Р. Шевельсоном метаанализ работ, опубликованных в период между 1997-м и 2007 гг. и посвященных вопросу о связи характеристик учителей и достижений учеников, показал, что в двух случаях из трех их авторы используют именно срезовые данные и корреляционный план. Известная проблема корреляционных исследовательских планов (чаще всего предполагающих регрессионный анализ полученных данных) состоит в том, отмечают Зидель и Шевельсон, что такой план не позволяет делать выводы о причинно-следственных связях между явлениями, не контролирует неизмеренные смещения и тем самым допускает множественное объяснение получаемых результатов. Именно экспериментальные или квазиэкспериментальные<sup>1</sup> подходы более обоснованы с точки зрения статистики, и их результаты часто показывают больший размер эффекта [1]. Общий вывод, который делают эти и другие авторы [1; 2], состоит в том, что подходы к изучению связей между различными характеристиками должны быть пересмотрены, и здесь необходимы именно экспериментальные дизайны.

Тем не менее, до сих пор экспериментальные или лонгитюдные данные назвать массовым явлением никак нельзя, особенно в России. Основная и часто блокирующая возможность проведения исследований с применением экспериментальных данных трудность состоит в том, что их сложно реализовать по целому ряду причин – этическим, финансовым, ресурсным [3, р. 1–2; 4, р. 608;

---

<sup>1</sup> Квазиэксперимент – модификация экспериментального плана для нелабораторных (естественных) ситуаций, когда исследователь не может манипулировать независимыми переменными (т.е. случайно приписывать какую-то комбинацию их значений каждому наблюдению) [6, р. 12; 7, с. 107–108].

5, p. 34; 6; 7]. В настоящее время компромиссом между наличием в основном срезовых данных и необходимостью изучать эффект воздействия одних переменных на другие признаются квазиэкспериментальные планы исследования, когда фактически *a posteriori* происходит переход от неэкспериментального плана исследования к экспериментальному. Подробно виды, особенности и недостатки квазиэкспериментальных планов будут рассмотрены ниже.

Мы предлагаем подходить к изучению связи характеристик учителей и достижений учеников в России с применением одного из методов построения регрессионных моделей с фиксированными эффектами<sup>1</sup>, а именно оценок первой разницы (ПР, *first difference estimators*), которые, при выполнении определенных условий, позволяют рассматривать данные срезовых исследований как квазиэкспериментальные. Статья имеет следующую структуру: обсуждение корреляционного и квазиэкспериментального планов исследований, описание примененного нами метода первой разницы для оценки регрессионных коэффициентов, описание проведенного исследования, его результаты, а также сравнение их с результатами, полученными путем регрессионных оценок с помощью метода наименьших квадратов (МНК).

## *Методология*<sup>2</sup>

Одно из самых очевидных отличий срезового дизайна от экспериментального заключается в том, что при эксперименте

---

<sup>1</sup> Модели с фиксированными эффектами – регрессионные модели, позволяющие контролировать данные по всем неизменяемым характеристикам респондентов, тем самым исключая возможность смещения получаемых результатов. Данные модели применимы в ситуациях, когда есть два и более измерения зависимой переменной [3].

<sup>2</sup> Мы не ставили здесь целью детально описать исследовательские планы. Мы лишь представим основные возможности и ограничения упомянутых выше подходов, релевантные теме нашего исследования.

изучается очень ограниченное (часто – один) количество факторов, которые могут быть источниками влияния на зависимую переменную, а возможные источники смешивающих влияний тщательно контролируются. Применение рандомизации обеспечивает случайное попадание участника в ту или иную группу. В срезовых же исследованиях задача учета именно одной характеристики обычно не ставится, и выборки, с одной стороны, перемешаны по большому количеству факторов, а с другой – могут быть смещены по какому-то из них. Выборки, как правило, выравниваются только с учетом половозрастной структуры генеральной совокупности. В связи с этим неправомерно делать выводы о причинно-следственной связи на основе корреляционных планов исследования на срезовых данных, речь идет только о статистической зависимости. К важным ограничениям для установления причинно-следственных связей можно отнести: смещенность выборки по значимым для исследования характеристикам, наличие неизмеренных смешивающих переменных (*confounding variable*), которые одновременно связаны как с зависимой, так и с независимыми переменными и др. (подробнее см.: [2; 3; 5; 8; 9]).

Распространен подход к решению этой проблемы, когда переменные, являющиеся потенциальными источниками смещения, включаются в регрессионные модели; тем самым контролируется или учитывается влияние этих переменных. Однако переменные, вносящие смещения, могут быть неизмеренными или даже не измеряемыми с достаточной надежностью<sup>1</sup>. Это множество всегда останется ограниченным измеренными переменными и никогда не может стать полным. Таким образом, регрессионные модели, построенные на срезовых данных, всегда допускают возможность смещения. Его величина может быть уменьшена, но вероятность самого смещения никогда не может быть исключена полностью.

---

<sup>1</sup> В качестве примера можно привести измерение экономического статуса семьи в результате опроса детей.

Если рассматривать центральную для исследования, которое мы описываем здесь, область связи определенных характеристик учителей, с одной стороны, и достижений их учеников – с другой, то попытки установить влияние первых на вторые ограничены возможностью неслучайного распределения учеников по учителям с разной степенью выраженности этих характеристик. Нередки ситуации, когда учителя неслучайным образом распределяются по школам: более квалифицированные учителя часто работают в более престижных и сильных школах. В то же время семьи с высоким социально-экономическим статусом (и, как правило, более успевающими детьми) отдают своих отпрысков в те же престижные школы. В таком случае анализ срезовых данных без учета возможных источников смещения, без сомнения, приведет к сдвигу результатов о связи изучаемых характеристик учителей и достижений учеников в положительную сторону. В реальности этой связи или нет вовсе, или она не столь сильна, однако большая часть наблюдаемого эффекта все же должна быть отнесена к неравномерному и неслучайному распределению учеников и учителей. Возможно обратное предположение: школы могут использовать своеобразную компенсаторную стратегию, когда более опытные учителя преподают в классах с более низкими достижениями учеников. Тогда в корреляционном дизайне те же характеристики будут иметь отрицательную (или положительную, но меньшую по модулю, чем в реальности) связь с достижениями учащихся. К тому же при наличии вступительных экзаменов в школы дети «выравниваются» по успеваемости. Таким образом, в регрессионных коэффициентах в характеристики учителей закладывается не только «эффект учителя», но и «эффект распределения учителей и учеников по школам» [10, р. 34–35; 11, р. 8].

Один из возможных подходов к решению этой проблемы – это, если позволяют данные, применить дополнительные процедуры, например, переход к таким регрессионным моделям, которые позволяют рассматривать исследовательский дизайн как квази-

экспериментальный. В литературе описано достаточно много таких подходов, из чаще применяемых можно назвать метод инструментальных переменных (*Instrumental variables*), оценку вероятности попадания под экспериментальное воздействие (*Propensity scores matching*), модели с фиксированными эффектами (*Fixed effects method*) и оценку прерывности регрессионной прямой в области появления экспериментального воздействия (*Regression discontinuity design*). Все эти методы нацелены на то, чтобы снизить эффект внешних неизмеренных переменных, которые могут быть неучтенной причиной изменений в результатах. Выбор того или иного метода определяется в основном особенностями данных, используемых при анализе<sup>1</sup>.

В рамках международных мониторинговых исследований школьного образования в основном применяются следующие квазиэкспериментальные планы. 1. Сравнение различий между когортами (*TIMSS*<sup>2</sup>: сравнение 4 и 8 классов). 2. Сравнение исторических трендов (*TIMSS*: сравнение результатов отдельно 4 или 8 классов за разные годы). 3. Изолирование возможных каузальных механизмов (модели с фиксированными эффектами). 4. Оценка прироста достижений за фиксированный период (*PISA: Regression Discontinuity*). 5. Сравнение успехов школьников внутри одной страны, которые при этом учатся по разным образовательным программам (сравнение частных и государственных школ: *Propensity Scores Matching*) [2, с. 284].

### *Модели с фиксированными эффектами*

Далее мы подробно рассмотрим метод оценки параметров регрессионных уравнений, называемый «Оценки первой разницы»

---

<sup>1</sup> Подробнее о квазиэкспериментальных планах исследования см., например, в [3; 4; 6] (мы указали наиболее поздние работы авторов).

<sup>2</sup> *Trends in International Mathematics and Science Study*. Международное исследование по оценке качества математического и естественно-научного образования. Подробнее об этом см. раздел «Данные».

(*First Difference Estimators*) в рамках регрессионных моделей с фиксированными эффектами.

Крупномасштабные исследования часто включают несколько разнесенных по времени повторных наблюдений за одним и тем же респондентом. Однако здесь может возникнуть проблема смещения по ненаблюдаемым характеристикам, коррелирующим как с зависимыми, так и с независимыми переменными. Для ее решения в рамках квазиэкспериментальных планов исследования контрольную и экспериментальную группы разделяют так, чтобы данная ненаблюдаемая характеристика оказалась в одинаковой степени присуща наблюдениям из обеих групп и таким образом была бы зафиксирована. При таком подходе основная сложность состоит в том, что необходимо строгое обоснование, какую именно характеристику фиксировать. При наличии двух наблюдений в разное время особенно сложно обосновать, что характеристика инвариантна по времени [5, р. 45–46].

Для данных *TIMSS*, на которых построено взятое нами для примера исследование, наиболее подходит один из вариантов метода фиксированных эффектов – метод оценок первой разницы (ПР) (*first difference*), поскольку в этом исследовании доступны результаты теста *TIMSS* по математике и ряду естественно-научных предметов каждого отдельного ученика и характеристики учителей, у которых он обучается этим же предметам. Следовательно, есть возможность, фиксируя личные характеристики ученика, который находится под педагогическим воздействием нескольких учителей, исключить из рассмотрения ненаблюдаемые переменные – неизмеренные характеристики учащихся, и минимизировать их влияние на результаты. Теперь все различия, которые демонстрирует ребенок в разных предметных областях *TIMSS*, могут быть приписаны различиям между его учителями. Закономерное предположение о разных способностях, которые может показывать ребенок в разных областях, например в математике и в биологии, может и будет исследовано дополнительно.

В литературе нет устоявшегося термина для обозначения такого рода квазиэкспериментов. По классификации, предложенной Ч. Рейхардтом [12], наш план исследования можно назвать «квазиэксперимент внутри одного респондента» (*Within-Participant Quasi-Experiments*). Имеется в виду, что в контрольной и экспериментальной группах участвуют, условно говоря, одни и те же респонденты. Следовательно, сравнение происходит не между двумя группами респондентов, подвергнувшихся различным воздействиям, а внутри одной группы респондентов, оказавшихся под двумя (одно из них может быть нулевым) разными воздействиями. Более узко такую методологию можно определить как «сравнение между предметами внутри одного ученика» (*Within-Student Between-Subject Approach*).

### *Оценки первой разницы*

В рамках моделей регрессий с фиксированными эффектами мы использовали метод оценок первой разницы (*first-difference estimators*) – ПР [11; 13; 14]. Оценки ПР для регрессий с фиксированными эффектами распространены в панельных исследованиях, где сравниваются характеристики одного и того же респондента в два момента времени. В нашем случае мы применяем оценки ПР не для двух моментов времени, а для одномоментного измерения двух характеристик (например, успеваемости по математике и физике) одних и тех же школьников.

Применительно к нашему исследованию основная идея, которая стоит за методом ПР, заключается в том, что мы рассматриваем каждого школьника как некое наблюдение, на которое оказывается экспериментальное воздействие, – он учится у двух учителей с разными характеристиками, а все остальные его свойства остаются зафиксированными. Иными словами, мы контролируем все индивидуальные характеристики, кроме одной – учитель, в чем, собственно, и заключается главное преимущество этого подхода.



Так как мы рассматриваем одного и того же ученика, можно предположить, что все его личные характеристики неизменны и одинаково сказываются на его успеваемости по различным школьным предметам. Таким образом, мы переходим к квазиэкспериментальному дизайну, который позволяет устанавливать причинные связи между характеристиками. Правда, расплатой за это окажется не слишком удобная интерпретация результатов, поскольку зависимая переменная – это теперь не просто балл ученика в тесте *TIMSS*, а разница между его баллами в двух тестах, причем каждый из тестов отвечает одному из уровней фактора, воздействующего на объект исследования. Оригинальное использование моделей с фиксированными эффектами при анализе панельных данных лишено этого недостатка, так как там учитывается разница между периодом времени  $t$  и предыдущим  $t - 1$ , и такая интерпретация как прирост/снижение значения очевиднее, чем увеличение разрыва в пользу одного предмета по сравнению с другим. Однако, по нашему мнению, преимущество построения квазиэкспериментального дизайна на основе срезовых (однократных) данных существенно превышает названные выше недостатки и значительно расширяет возможности работы с данными международных сравнительных исследований.

### *Оценка регрессионных коэффициентов*

Как уже говорилось выше, находясь в рамках исследования *TIMSS*, мы имеем возможность рассматривать достижения каждого ученика как функцию от его собственных характеристик, характеристик учителя, который ведет тот или иной предмет, и профессиональных практик, которые он применяет. Таким образом мы рассчитали серию регрессионных уравнений (МНК):

$$Y_{Mij} = \beta_0 + \beta_1 P_{ij} + \beta_2 T_{Mj} + \beta_3 C_j + \varepsilon_{Mij} \quad (1)$$

$$Y_{\Phi ij} = \beta_0 + \beta_1 P_{ij} + \beta_2 T_{\Phi j} + \beta_3 C_j + \varepsilon_{\Phi ij} \quad (2)$$

где  $\beta_0$  – константа;  $Y_{Mij}$  ( $Y_{Фij}$ ) – результат теста TIMSS по математике (физике)  $i$ -го ученика  $j$ -й школы<sup>1</sup>;  $P_{ij}$  – вектор характеристик ученика;  $T_{Mj}$  – вектор характеристик учителя;  $C_j$  – вектор характеристик учебного процесса;  $\varepsilon_{Mij}$  – ошибка.

Подобные уравнения были рассчитаны для всех четырех изучаемых нами предметов.

Далее мы перешли к оценке моделей ПР. Путем вычитания уравнения 2 из уравнения 1, можно уйти от присутствия в модели индивидуальных характеристик учеников, тем самым зафиксировав их:

$$\Delta Y_i = \beta_0 + \beta_2 \Delta T_j + \beta_3 \Delta C_j + \Delta \varepsilon_{ij}, \quad (3)$$

где  $\Delta$  означает разницу между двумя значениями в первом и втором наблюдениях; остальные обозначения те же, что и в уравнениях 1 и 2.

Далее оценки регрессионных коэффициентов получаются с помощью МНК регрессии [13, с. 6]. Таким образом, мы избавились от одного из недостатков корреляционных исследований – невозможности учесть все характеристики школьников, которые могут быть связаны как с их достижениями, так и с характеристиками учителей, у которых они учатся.

Обычно при применении ПР оценок делается предположение [13; 15; 16], что регрессионные коэффициенты  $\beta_2$  и  $\beta_3$  равны для различных предметов (или что регрессионные линии для уравнений 1 и 2 параллельны), что означает, что та или иная характеристика учителя или педагогическая практика окажет одинаковый эффект на достижения учеников вне зависимости от предмета. Однако нет оснований предполагать, что это так, как и нет оснований для обратного вывода<sup>2</sup>.

---

<sup>1</sup> Так как в выборку TIMSS попадает только один класс в каждой школе, то уровни класса и школы в данном случае совпадают.

<sup>2</sup> Впрочем, косвенным обоснованием здесь может служить высокая корреляция результатов по всем предметам.

Таким образом,  $\Delta$  в случае характеристик учителей и их педагогических практик может подразумевать две ситуации: 1) значения переменных действительно вычитаются для двух разных наблюдений, например, из стажа учителя по математике вычитается стаж учителя по физике, и в окончательное уравнение, попадает одна переменная; 2) фактически в уравнение попадают две переменные для каждого учителя.

Уравнение для первой ситуации:

$$\Delta Y_i = \beta_0 + \beta_2(T_{Mj} - T_{Фj}) + \beta_3(C_{Mj} - C_{Фj}) + \Delta \varepsilon_{ij}. \quad (4)$$

Для второй:

$$\Delta Y_i = \beta_0 + (\beta_2 T_{Mj} - \eta_1 T_{Фj}) + (\beta_3 C_{Mj} - \eta_2 C_{Фj}) + \Delta \varepsilon_{ij}. \quad (5)$$

Однако в связи с тем, что часть наших переменных дихотомические, и вычитать одно из другого не имеет смысла, мы использовали либо уравнения типа (4), либо смешанные модели, в которых переменные «пол», «стаж», «образование учителя», «объем домашних заданий» включены в модель отдельно для учителя математики и отдельно для учителя одного из естественно-научных предметов, а для переменных, характеризующих педагогические практики и время преподавания предмета в неделю, вычислялась разница.

В зависимости от того, какой набор независимых переменных включался в уравнение, мы выделили минимальную и полную модели (и в МНК, и в ПР моделях). В минимальную модель входили только характеристики учителей, в полную – характеристики и учителей, и учебного процесса<sup>1</sup>.

### Особенности интерпретации регрессионных коэффициентов в уравнениях ПР

Так как интерпретация уравнений с разностями в левой части несколько затруднительна, опишем здесь основные ее моменты.

---

<sup>1</sup> В связи с мультиколлинеарностью переменные включались в модель по очереди.

Первое: формально возможна только относительная интерпретация роста или снижения баллов по одному предмету по сравнению с баллами по другому предмету. Иными словами, сравнивая, например, математику и физику, мы можем сказать только, что применение той или иной педагогической практики увеличивает или сокращает разницу в результатах тестов по двум предметам. Но, поскольку независимые переменные относятся либо к учителю математики, либо к учителю физики, то это вносит ясность в интерпретацию – повышаются/снижаются ли баллы по одному из предметов. Иными словами, рассматривая одну из независимых переменных, мы каждый раз можем делать вывод об одном из предметов.

Второе: об интерпретации знаков регрессионных коэффициентов. Поскольку зависимая переменная – это разница между двумя баллами (например, математика – физика), то направление связи с независимыми переменными должно интерпретироваться по-разному. Стоит обратить внимание, что связи интерпретируются только внутри одного предмета и только в том случае, если оценивается уравнение 5. При уравнении 4 интерпретироваться может только величина коэффициента без направления связи.

Коэффициенты при переменных, кодирующих учителей математики (или для предмета слева от знака минус в зависимой переменной):

– положительный коэффициент при признаке учителя математики увеличивает разницу между баллами по математике (М) и баллами по физике (Ф), следовательно, успеваемость по М растет.

– отрицательный коэффициент при признаке учителя М сокращает разницу между баллами по М и баллами по Ф, следовательно, успеваемость по М падает.

Коэффициенты при учителях физики (или для предмета справа от знака минус в зависимой переменной):

– положительный коэффициент при характеристике учителя физики увеличивает разницу между баллами по М и баллами по Ф, следовательно, результаты по Ф падают.

– отрицательный коэффициент при признаке учителя физики сокращает разницу между баллами по М и баллами по Ф, следовательно, успеваемость по Ф растет.

### Кластерная коррекция (cluster correction)

Данные TIMSS имеют иерархическую структуру [17]. Опрашивая по анкетам и учителей и учеников, исследователи закладывают в выборку три уровня: класс (опрашиваются все ученики в классе) → учителя → школа. Таким образом, мы получаем данные от группы учеников, обучающихся у одного учителя (или группы учителей по разным предметам). Следовательно, наблюдения на уровне класса нельзя назвать независимыми друг от друга. Это приводит к заниженной оценке стандартных ошибок коэффициентов регрессии. Процедура кластерной коррекции позволяет учесть этот факт и рассчитать корректные ошибки.

## *Данные*

Начиная с конца XX в. в сфере социальных наук и образования проводится все больше крупномасштабных международных исследований, причем Россия принимает участие в подавляющем большинстве из них. Среди исследований школьного образования самые известные – уже упоминавшийся выше TIMSS, Изучение качества и понимания текста (*Progress in International Reading Literacy Study (PIRLS)*) и Международная программа по оценке качества образовательных достижений учащихся (*Programme for International Student Assessment (PISA)*). В настоящее время все эти исследования оказались в сфере внимания лиц, принимающих решения в области образовательной политики, и интерес к ним растет, так как они потенциально содержат в себе информацию по широкому кругу вопросов. Мы остановились на изучении данных одного из них: а именно TIMSS. В исследовании TIMSS опрашиваются следующие когорты: ученики (тест по 4 предметам:

математика, физика, химия и биология<sup>1</sup>, а также анкета с контекстной информацией); учителя, ведущие предметы, по которым проводится тестирование; представитель администрации школы. Опрос проводится в 4 и 8 классе.

Наше исследование строилось на российских данных *TIMSS* 2007 г., 8 класс. Кроме результатов тестирования школьников, мы использовали данные из анкеты для учителей. Такая комбинация позволила нам сопоставить результаты учеников по 4 предметным тестам с контекстными данными об учителях, ведущих эти предметы.

Российская выборка *TIMSS* 8 класс в 2007 г. состояла из 4472 учащихся из 210 общеобразовательных учреждений 58 регионов страны. Выборка репрезентативна по школам, в каждой школе опрашивается целый класс (с учетом доступности детей для опроса в день проведения исследования) и учителя, преподающие в данном классе названные выше предметы.

Для построения модели, подбора независимых переменных, которые в теории могли бы быть связаны с достижениями учеников, из русскоязычной версии анкеты *TIMSS* мы выбрали все вопросы об образовании, стаже и тому подобных формальных характеристик учителей, а также переменные, описывающие различные практики преподавания на уроках. По возможности мы выбирали такие переменные, которые можно отнести к объективным, фактическим аспектам функционирования учителя в рамках учебного процесса, а не к самооценочной информации (ограничение связано с ее надежностью) [18].

Кроме того, мы включили в модель пол ученика, так как обнаружили статистически значимые различия между достижениями мальчиков и девочек. Предположив, что учителя с определенными характеристиками неравномерно распределяются по школам и, сле-

---

<sup>1</sup> Результаты тестов измеряются в шкале от 1 до 1000 баллов, но реальное распределение варьируется примерно от 200 до 800 баллов.

довательно, по ученикам с различным социально-экономическим статусом (СЭС), мы контролировали базовые МНК модели по показателям СЭС ученика. В модели ПР мы включали средний СЭС по классу как показатель эффекта совместного обучения (*peer-effect*).

В связи с особенностями формулировок вопросов об образовании учителя в анкете TIMSS, мы рассматривали не его фактический уровень или область, а широту образования учителя. Например, специализировался ли учитель, скажем, только по математике или еще по какому-либо другому предмету, или нет<sup>1</sup>. На основе вопроса «Каковы основные области Вашего профессионального образования?» были построены следующие категории<sup>2</sup>:

Учителя математики:

1. Только математика / математика И методика преподавания математики;

2. Математика И методика преподавания математики И общая педагогика/математика И методика преподавания математики И общая педагогика И естественные науки И/ИЛИ методика преподавания естественно-научных дисциплин;

3. Все остальные варианты.

Учителя физики:

1. Только физика / физика И методика преподавания естественно-научных дисциплин / физика И любая другая специализация И нет методики преподавания естественно-научных дисциплин И нет общей педагогики;

---

<sup>1</sup> Мы не рассматривали ответы на вопрос: есть ли у учителя, в принципе, образование в области преподаваемого предмета, так как оно есть практически у всех школьных учителей. То же самое можно сказать про уровень образования – разделение на бакалавриат и магистратуру еще нерелевантно для российской ситуации.

<sup>2</sup> Имеются в виду логические И (т. е. респондент отметил, что он изучал и ту и другую специализацию) и ИЛИ (часть респондентов в данной категории могли как изучать, так и не изучать данную специализацию. Знак «/» отделяет один тип образования учителя от другого. В укрупненные категории попадали два или три типа образования.

2. Физика И методика преподавания естественно-научных дисциплин И общая педагогика / физика И методика преподавания естественно-научных дисциплин И общая педагогика И/ИЛИ любая другая специализация;

3. Все остальные варианты.

Для учителей биологии и химии категории были аналогичны категориям для учителей физики.

Анкеты также включала вопросы о стаже учителя, количестве минут преподавания данного предмета в неделю, наконец, практике учителя: объем домашнего задания, распределение времени урока на разные виды деятельности (в %), их распространенность на уроке (по частоте). Ответы о распределении времени урока были сгруппированы в четыре категории: фронтальное обучение, индивидуальная работа, контроль усвоения, непродуктивно потраченное время. В вопросе о распространенности различных видов деятельности на уроке также было выделено четыре вида деятельности: репродуктивное обучение, задания на понимание и интерпретацию, задания на развитие метанавыков, групповая работа (признак измерялся от 0 до 1).

В качестве зависимой переменной выступала стандартизированная разница в результатах написания теста *TIMSS* по математике и естественно-научным дисциплинам. Для каждого ученика были рассчитаны три дополнительных переменные: 1) (балл по математике) – (балл по физике); 2) (балл по математике) – (балл по химии); 3) (балл по математике) – (балл по биологии). Затем для каждой из них были построены регрессионные модели. Независимые переменные, которые включались в модель, можно разбить на две группы: характеристики учителя (образование, стаж, пол) и его педагогические практики (время преподавания дисциплины, количество домашних заданий, педагогические практики на уроках). В ходе исследования оценивались три модели, включающие либо один из наборов переменных, либо полные модели, куда включались все переменные.



Обзор аналогичных исследований<sup>1</sup> показывает, что при использовании данных *TIMSS* наиболее часто рассматривались такие характеристики учителей, как образование (наличие диплома о высшем образовании, а именно педагогическом или в области преподаваемого предмета, уровне образования) [10; 15; 16; 19], а также стаж, педагогические приемы при проведении уроков и подготовке к ним, самооценка готовности вести тот или иной предмет, профессиональные установки и ожидания [10; 15; 16; 19; 20; 21; 22]. Помимо этого, рассматриваются социально-демографические характеристики [16; 19], а также численность класса, характеристики системы образования [23].

Несколько работ с использованием регрессионных моделей показали крайне рассогласованные результаты, иногда даже в рамках одного и того же исследования [10; 19]. В работах, позволяющих сравнить эффекты характеристик учителей между МНК и ПР моделями, показано, что ПР дают меньше значимых связей, однако больше согласованных и интерпретируемых результатов [10; 16; 22]. На основе наших данных получилась несколько иная картина, о чем будет рассказано ниже.

## *Результаты*

В связи с методологическим фокусом данной статьи мы опустим результаты описательной статистики, и сразу перейдем к рассмотрению регрессионных моделей. Отметим только, что, как мы говорили выше, применяя модели с фиксированными эффектами, мы делаем предположение о неизменности всех характеристик

---

<sup>1</sup> Выбор исследований был ограничен следующими критериями: 1) в качестве зависимой переменной выступали результаты теста *TIMSS* 8 класс по математике и естественным наукам); 2) независимыми переменными послужили также характеристики учителей, доступные в анкете *TIMSS*; 3) использовался регрессионный анализ (МНК или методом первой разницы).

школьника, включая его способности к различным предметам. Для проверки данного предположения мы вычислили коэффициенты корреляции Пирсона между баллами школьников по математике и остальным предметам. Результаты показали, что корреляция во всех случаях равна приблизительно 0,8 ( $p < 0,001$ ). Это позволяет нам говорить, что у учеников, склонных к изучению математики, наблюдается примерно такая же склонность к изучению естественно-научных дисциплин, и подтверждает адекватность применения регрессий с фиксированными эффектами.

Далее опишем полученные результаты по каждому предмету, параллельно сравнивая МНК и ПР оценки.

### Математика

Если рассматривать МНК модели, то статистически значимыми оказались четыре переменные, касающиеся учебного процесса, и не одной переменной, описывающей формальные характеристики учителя. Положительно связаны с результатами TIMSS по математике такие характеристики, как количество минут преподавания в неделю, что достаточно логично: чем больше времени ученики изучают предмет, тем больше материала они успевают освоить, даже если не происходит роста эффективности обучения. Есть вероятность, что здесь играет роль взаимодействие между количеством уроков математики в неделю и типом школы (про который нам ничего не известно), однако данная переменная оказалась статистически значима в моделях ПР, что исключает данное предположение. По крайней мере мы можем утверждать, что время преподавания играет роль в любой школе. Кроме этой характеристики, положительно на повышении баллов сказываются количество заданий на развитие метанавыков и объем индивидуальной работы. Связь других видов деятельности с количеством фронтального преподавания на уроке оказалась отрицательной. Оценки методом ПР, кроме уже названного количества минут

преподавания, подтвердили отрицательное влияние фронтального преподавания. Также со знаком минус стало статистически значимым репродуктивное обучение, что полностью укладывается и в логику МНК, и в ожидания с точки зрения здравого смысла.

Однако переменные «задачи на развитие метанавыков» и «индивидуальная работа» оказались статистически значимы только в МНК моделях, и это свидетельствует, что здесь присутствует смещение по каким-либо ненаблюдаемым переменным. Самое логичное предположение состоит в том, что смещение проходит по типу школы.

### Физика

Как и в ситуации с математикой, в МНК моделях большее количество переменных, связанных с педагогическими практиками, оказались статистически значимыми. Причем те переменные, которые отвечают за практики, считающиеся современными, «правильными», более эффективными, положительно связаны с достижениями. Но, как и в случае с математикой, здесь можно предположить смещение по типу школы. Иными словами, не упомянутые педагогические практики приводят к высоким результатам, а тот факт, что эти практики применяются в более статусных школах с более сильными учениками. Оценки ПР, позволяющие исключить характеристику ученика, в том числе и СЭС его семьи, из перечня влияющих факторов, показывают, что вывод о большей эффективности данных методов не обоснован. Напротив, ПР оценки доказывают положительный эффект репродуктивного обучения, а также женского пола учителя.

### Биология

И здесь складывается похожая ситуация – МНК модели показали значимость педагогических практик, однако в ПР оценках они не подтвердились. Но зато статистически значимыми оказа-

лись стаж работы (причем чем больше стаж, тем заметнее растут достижения ученика, например, в биологии по сравнению с математикой, т. е. ученик получает по биологии такие же оценки, как и по математике), пол учителя и объем домашнего задания.

### *Химия*

Как в биологии и математике, в МНК значимыми оказались только характеристики учебного процесса. Фронтальное обучение, как и в случае трех других дисциплин, связано с достижениями отрицательно, а понимание и развитие метанавыков у школьников, так же как и их индивидуальная работа – положительно. Однако при переходе к квазиэкспериментальным моделям картина совершенно меняется: ни одна из названных выше характеристик не остается статистически значимой в ПР модели. В данном случае обнаружилось, что положительное влияние на рост достижений оказывают следующие характеристики: женский пол учителя, опыт преподавания от 3 до 10 лет и, как ни странно, репродуктивное обучение (что в случае с химией предполагает заучивание материала наизусть, чтение учебника, наблюдение за экспериментами, которые проводит учитель, объяснение увиденного). Также по результатам ПР получается, что более широкое с точки зрения полученной специализации образование учителя негативно влияет на рост достижений его учеников.

### *Заклочение: сравнение результатов МНК и ПР оценок*

Как видно из описанных выше результатов, ПР и МНК оценки практически не совпадают. В МНК, в отличие от ПР, много статистически значимых характеристик, касающихся педагогических практик. Однако можно предположить, что педагогические практики связаны со сложившимися в школе традициями и,

соответственно, неравномерно распределены по школам. Причем выявилось больше именно ложноположительных результатов, т. е. чаще встречается такой вид смещения, когда более квалифицированные учителя работают в более престижных школах. Таким образом, результаты МНК регрессии подвержены смещениям по данным переменным, а ПР оценки позволяют это проконтролировать. В свою очередь, отсутствие противоречия между МНК и ПР результатами может трактоваться как подтверждение того факта, что определенная характеристика свойственна всем учителям в равной степени, вне зависимости, например, от социального статуса.

Таким образом, переход к квазиэкспериментальному плану исследования и применение для этих целей регрессий с фиксированными оценками позволил нам отделить связи, обусловленные корреляцией характеристик учителей и достижений учеников, от связей, обусловленных неравномерным распределением учеников по учителям и школам. Иными словами, отделить эффекты учителей от смешанных эффектов ученик – учитель – школа.

Резюмируя, можно заключить, что на основе МНК регрессий мы можем сделать выводы о связи между различными характеристиками, например, сказать, что более низкие результаты связаны с количеством фронтального обучения, а более высокие – с заданиями на метанавыки, но у нас нет оснований делать выводы о причинно-следственных связях. Так, мы не можем сказать, что фронтальное обучение снижает результаты, поскольку есть возможность альтернативного объяснения: когда учитель видит, что ученики слабые, он не стремится применять (или не может реализовать) никакие другие педагогические практики, кроме фронтального обучения. Кроме того, учителя, применяющие в образовательном процессе более разнообразные педагогические практики, могут работать (и скорее всего работают) в школах с более успевающими детьми, что опять ставит результаты впереди методов преподавания, т. е. причина и следствие меняются местами.

Сопоставление результатов ПР и МНК моделей может указать на те характеристики, по которым есть смещение и, следовательно, они отражают неслучайное распределение учеников по учителям или школам. В нашем случае это педагогические практики. Далее, зная это, мы можем предположить, какие неизмеренные в исследовании характеристики могут быть связаны с данными переменными и с достижениями учеников, т. е. выступать смешивающими переменными (*confounding variable*). Например, статус школы.

Каковы ограничения применения метода ПР? Во-первых, для его применения требуются специфические данные – наличие двух ситуаций «воздействия» на одного респондента. Во-вторых, делая выводы, все, о чем мы можем судить, это то, что результаты улучшаются или ухудшаются по одной дисциплине относительно другой. Однако здесь, на наш взгляд, минусы не перевешивают названные выше плюсы. И метод, безусловно, заслуживает внимания исследователей в области социальных наук.

#### ЛИТЕРАТУРА

1. *Seidel T., Shavelson R.J.* Teaching Effectiveness Research in the Past Decade: the Role of Theory and Research Design in Disentangling Meta-analysis Results // Review of Educational Research. 2007. Vol. 77. No. 4. P. 454–499.
2. *Raudenbush S.W., Kim J.S.* Statistical Issues in Analysis of International Comparisons of Educational Achievement // Methodological Advances in Cross-national Surveys of Education Achievement / Ed. A.C. Porter, A. Gamoran. Washington, DC: National Academies Press, 2002.
3. *Allison P.D.* Fixed Effects Regression Methods for Longitudinal Data Using SAS. Cary, NC: SAS Institute, 2005.
4. *Shadish W.R., Cook T.D.* The Renaissance of Field Experimentation in Evaluating Interventions // Annual Review of Psychology. 2009. Vol. 60. P. 607–629.
5. *Schneider B., Carnoy M., Kilpatrick J., Schmidt W.H., Shavelson R.J.* Estimating Causal Effects Using Experimental and Observational Designs (Report from the Governing Board of the American Educational Research Association Grants Program). Washington, DC: American Educational Research Association, 2007.
6. *Shadish W.R., Cook T.D., Campbell D.T.* Experimental and Quasi-experimental Designs for Generalized Causal Inference. Boston: Houghton Mifflin Company, 2002.
7. *Кэмпбелл Д.* Модели экспериментов в социальной психологии и прикладных исследованиях. М.: Прогресс, 1980.

8. *Freedman D.A.* Statistical Models. Cambridge: Cambridge University Press, 2009.
9. *Raudenbush S.W.* Learning from Attempts to Improve Schooling: The Contribution of Methodological Diversity // Educational Researcher. 2005. Vol. 34. No. 5. P. 25–31.
10. *Pumsaran T.* Teacher's Educational Background and Student Achievement in Thailand: A Dissertation Submitted to the School of Education and the Committee on Graduate Studies of Stanford University. 2010. Ch. 2.
11. *Clotfelter C., Ladd H.F., Vigdor J.L.* Teacher Credentials and Student Achievement in High School: A Cross-subject Analysis with Student Fixed Effects. NBER Working Papers 13617. 2007.
12. *Reichardt C.S.* Treatment Effects // Encyclopedia of Social Measurement. N.Y.: Elsevier, 2005. P. 875–882.
13. *Dee T.S., Cohodes S.R.* Out-of-field Teachers and Student Achievement: Evidence from Matched-pairs Comparisons // Public Finance Review. 2008. Vol. 36. P. 7–32.
14. *Brüderl J.* Panel Data Analysis. University of Mannheim, March 2005 [online]. URL: <http://www2.sowi.uni-mannheim.de/lsssm/veranst/Panelanalyse.pdf> (дата обращения: 01.10.2012).
15. *Schwerdt G., Wuppermann A.C.* Is Traditional Teaching Really All That Bad? A Within-student Between-subject Approach. CESIFO Working Paper No. 2634. Category 5: Economics of Education. 2009.
16. *Klaveren C., van.* Lecturing Style Teaching and Student Performance // Economics of Education Review. 2011. Vol. 30. No. 4. P. 729–739.
17. TIMSS 2007 Technical Report / Ed. J.F. Olson, M.O. Martin, I.V.S. Mullis. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College, 2008. Ch. 5.
18. *Тюменева Ю.А., Хавенсон Т.Е.* Международные сравнительные исследования качества образования для оценки эффективности реформ: проблемы, ограничения и возможные решения // Материалы 8-й международной научно-практической конференции. М.: Дело, 2012.
19. *Zuzovsky R.* Teachers' Qualifications and Their Impact on Student Achievement: Findings from TIMSS 2003 Data for Israel // IERI Monograph Series. Issues and Methodologies in Large-Scale Assessments. 2009. Vol. 2. P. 37–62.
20. *Yee K.M.* Head of the Class: Exploring the Link between Teacher Quality, Instructional Practice, and Student Outcomes in Indonesia, Malaysia, and the Philippines: A Thesis Submitted to the Faculty of the Graduate School of Arts and Sciences of Georgetown University. Washington, D.C., 2010.
21. Contexts of Learning Mathematics and Science: Lessons Learned from TIMSS / Ed. S.J. Howie, T. Plomp. L.: Routledge, 2006.

22. *Falch T., Rønning M.* Homework Assignment and Student Achievement in OECD Countries. Working Paper Series, Norwegian University of Science and Technology. 2011. No. 5.

23. *Pong S., Pallas A.* Class Size and Eighth-grade Math Achievement in the United States and Abroad // Educational Evaluation and Policy Analysis. 2001. Vol. 23. No. 3. P. 251–273.