
ПРАКТИКИ СБОРА И АНАЛИЗА ФОРМАЛИЗОВАННЫХ ДАННЫХ

О.В. Волченко, А.А. Широканова
(Санкт-Петербург)

ПРИМЕНЕНИЕ МНОГОУРОВНЕВОГО РЕГРЕССИОННОГО МОДЕЛИРОВАНИЯ К МЕЖСТРАНОВЫМ ДАННЫМ (НА ПРИМЕРЕ ГЕНЕРАЛИЗОВАННОГО ДОВЕРИЯ)¹

В фокусе данной статьи – метод многоуровневого регрессионного моделирования: обосновываются его преимущества перед регрессией методом наименьших квадратов в сравнительных исследованиях с иерархической структурой данных. Представлены содержательные и статистические критерии применения многоуровневого моделирования, дан обзор реализации метода в различных пакетах статистического анализа. На межстрановых данных из базы данных Всемирного исследования ценностей пошагово построена и протестирована двухуровневая линейная модель связи генерализованного доверия и уровня образования («нулевая

Олеся Викторовна Волченко – аспирант, Национальный исследовательский университет «Высшая школа экономики» – Санкт-Петербург; стажер-исследователь, Лаборатория сравнительных социальных исследований, Национальный исследовательский университет «Высшая школа экономики». E-mail: ovolchenko@hse.ru.
Анна Александровна Широканова – кандидат социологических наук, старший научный сотрудник, Лаборатория сравнительных социальных исследований, Национальный исследовательский университет «Высшая школа экономики»; доцент, Департамент социологии, Национальный исследовательский университет «Высшая школа экономики» – Санкт-Петербург. E-mail: ashirokanova@hse.ru.

¹ Исследование финансировалось в рамках государственной поддержки ведущих университетов Российской Федерации «5-100». Работа выполнена в рамках гранта РФФИ № 16-36-00372 «Формальный, неформальный и семейный социальный капитал: кросс-культурный анализ с использованием иерархического регрессионного моделирования».

модель», модель с фиксированным интерсептом, модель со случайным интерсептом, модель со случайными эффектами и интерсептом, модель с включением предиктора второго уровня, модель с межуровневым взаимодействием). Описаны меры диагностики качества многоуровневой модели (AIC, BIC, функции максимального правдоподобия, псевдо- R^2) и способы проверки устойчивости полученных результатов. В последней части представлен обзор применяемых в социальных науках альтернатив многоуровневому моделированию иерархических данных. Приложен код для репликации результатов в статистической среде *R*.

Ключевые слова: сравнительные исследования, иерархические модели, многоуровневое моделирование, WVS, Всемирное исследование ценностей, доверие, статистическая среда *R*, регрессионный анализ, межстрановые данные.

Введение: иерархические структуры в данных

В социальных науках, особенно в психологии, исследованиях образования и социологии, многоуровневое¹ моделирование постепенно превращается в новый стандарт анализа иерархических данных. Не случайно именно этой теме был посвящен один из последних выпусков журнала сравнительных исследований «European Sociological Review» [3, p. 1–2], а методологи пишут о «вездесущности иерархического линейного моделирования» [4]. Хотя в русскоязычных публикациях при помощи многоуровневых регрессионных моделей уже анализировались межстрановые

¹ В различных дисциплинах для названия метода используются термины «многоуровневая модель» (социология), «модель со смешанными эффектами», «смешанная модель» и «модель со случайными (random) эффектами» (биометрика), «регрессионная модель со случайными коэффициентами» (эконометрика), «модели с компонентами ковариации» (статистика) [1, p. 5–6]. Ряд авторов считают наиболее содержательным термин «иерархические модели» (например: [1, p. 6; 2, p. 2]).

данные [5; 6; 7; 8], региональные данные [9; 10], данные о школах [11; 12; 13] и результаты когнитивных исследований [14; 15], методологических публикаций о многоуровневом моделировании на русском языке до сих пор не было. Цель данной статьи – раскрыть логику и процедуру построения многоуровневых линейных моделей и продемонстрировать их использование на примере индивидуальных и страновых факторов обобщенного доверия.

Многие экспериментальные и опросные данные имеют естественную иерархическую структуру, когда наблюдения не независимы друг от друга, а «вложены» в более крупные географические и организационные единицы¹, будь то школа и район – для учеников школ, страна и регион – для жителей, организация – для работников и т.д. [1, р. 3–4; 16, р. 8; 17, р. 9].

Как отмечают статистики И. Крефт и Я. де Лёв, понимание иерархической структуры данных изменяет видение окружающего мира, заставляя повсюду отмечать разные уровни наблюдений [18, р. 1]. Хотя представление об иерархичности данных в том или ином виде задавалось и ранее – при анализе панельных данных, введении модерирующих контекстуальных переменных и др., – однако именно многоуровневое моделирование предоставило удобный статистический инструмент для анализа иерархических данных, позволяя учитывать эффекты² разноуровневых переменных, а также – при соблюдении предположений метода – в рамках одной модели давать статистически надежные ответы на вопрос: как характеристики макроединиц влияют на отношения на микроуровне?

¹ Единицы второго и последующих уровней в литературе называют классами, группами, контекстами, кластерами и т. д.

² Использование термина «эффект» в словосочетаниях «фиксированный эффект», «случайный эффект», «эффект межуровневого взаимодействия» в многоуровневом регрессионном моделировании является традиционным и здесь и далее не обязательно предполагает наличие причинно-следственных отношений между переменными.

Последствия ошибок в моделировании структуры иерархических данных наиболее опасны в исследованиях, на основе которых принимаются управленческие решения. Так, в известном британском исследовании 1976 г. Н. Беннетом было установлено, что формальный стиль преподавания обеспечивает более высокие экзаменационные оценки по сравнению со смешанным и неформальными стилями преподавания. Однако его модель не учитывала, что опрашивались ученики 36 преподавателей из разных школ, а данные собирались по трем предметам в начале и в конце года. Повторный анализ данных с учетом их иерархической структуры и автокорреляции остатков по группам (преподаватель, предмет, время) показал, что формальный стиль работает лучше неформального только в преподавании языка; во всех остальных случаях различия были не значимы [19, р. 48]. Со статистической точки зрения, даже при самом простом, срезовом дизайне исследования, если наблюдается автокорреляция остатков внутри группы (например, ученики одного учителя лучше усваивают одни темы и хуже – другие, но темы различаются между учителями), ошибки коэффициентов на уровне группы (учителя, района и т.д.) в регрессионной модели методом наименьших квадратов (МНК) будут оценены неправомочно оптимистично, что приведет к росту ложноположительных результатов. В свою очередь, учет иерархической структуры данных в многоуровневой модели позволяет не только должным образом идентифицировать эффекты групповых единиц наблюдения (это возможно и при других методах), но и моделировать эти эффекты.

Почему обычная МНК-регрессия дает ложноположительные результаты на иерархических данных? В регрессионных моделях данных с многоуровневой структурой, где используется метод наименьших квадратов, иерархия насильственно уплощается: так, в межстрановых исследованиях данные странового уровня представляются как общая точка пересечения с осью ординат

при нулевом значении всех предикторов¹ для всех представителей данной страны. Однако целью исследования может быть не просто зафиксировать необъясненную дисперсию по каждой группе, а определить зависимость между групповыми и индивидуальными переменными вне связи с конкретной страной или же установить, при каких страновых условиях наблюдается связь между переменными на индивидуальном уровне. В таком случае регрессия методом наименьших квадратов будет ошибочным выбором, поскольку нарушается одно из важных предположений МНК-регрессии – о независимости наблюдений – и в результате она дает слишком оптимистичные оценки стандартных ошибок коэффициентов регрессии (ложноположительная значимость эффектов вследствие инфляции ошибки I рода) [20, р. 196]. Применение робастных кластеризованных остатков, распространенное в эконометрике, дает несмещенные ошибки коэффициентов, но с содержательной точки зрения решает задачу лишь отчасти (см. ниже) [21, р. 5–6].

На *рис. 1* обобщенно изображены особенности данных, которые могут привести к неверным выводам при применении МНК-регрессии. *Рис. 1а* иллюстрирует случай экологической ошибки (ecological fallacy), или эффекта Робинсона, когда взаимосвязи между переменными на индивидуальном и групповом уровне различаются, но при этом интерпретации связей на групповом уровне применяются к индивидуальным объектам [17, р. 15; 22, р. 343]. Например, если в странах с более высоким средним доходом уровень доверия незнакомцам выше, из этого не следует, что индивиды с более высоким уровнем дохода доверяют другим сильнее, чем индивиды с низким доходом. Распространение вывода о связи дохода и доверия, полученное на основе страновых

¹ Многоуровневое моделирование предполагает возможность вариации точек пересечения регрессионной прямой оси Y («модель случайных интерсептов» в литературе по многоуровневому моделированию), поэтому вместо термина «свободный член» здесь и далее используется «интерсепт».

данных, на индивидуальный уровень в данном случае привело бы к экологической ошибке.

Другой случай (*рис. 1б*) – взаимная компенсация положительных и отрицательных эффектов в разных группах [17, р. 58]. Например, в некоторых странах более высокое образование связано с повышением доверия незнакомым людям, а в других – наоборот. Среднее значение эффекта в таком случае может быть близко к нулю, однако оно не будет отражать разнообразия наблюдаемых эффектов, и сам эффект, таким образом, может быть не замечен.

Третья возможная проблема – коррелирующие ошибки (*рис. 1в*), когда значение зависимой переменной объясняется принадлежностью к группе, а уже принадлежность к конкретной группе определяется независимой переменной [23]. Например, частота религиозных практик может варьироваться между религиями, но быть однородной внутри религии; если религия индивида предписывается этничностью, тогда МНК-регрессия приведет к обнаружению ложной связи между этничностью и частотой религиозных практик, тогда как они могут быть связаны лишь опосредованно, через религию.

Дальнейшая структура статьи такова: в первой части описана многоуровневая модель; во второй – представлены содержательные и статистические критерии целесообразности ее применения. В третьей части пошагово описан процесс оценки двухуровневой модели, который далее рассмотрен на примере генерализованного доверия (на межстрановых данных Всемирного исследования ценностей, WVS [24]). Представлены способы проверки надежности результатов. В последней части дается обзор альтернатив многоуровневому моделированию при работе с данными, обладающими иерархической структурой.

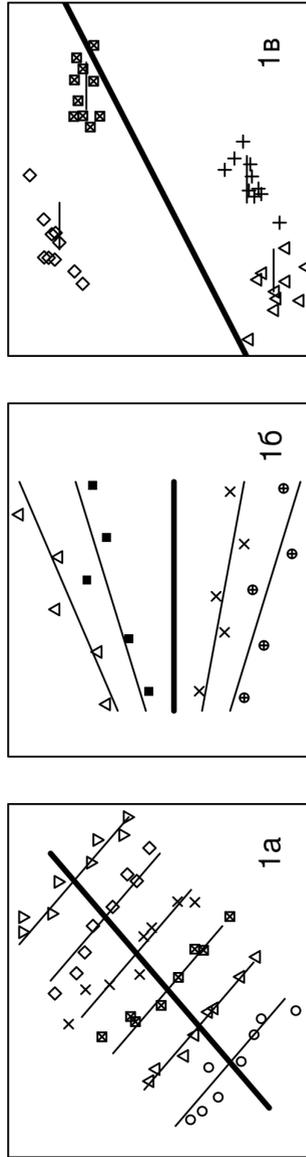


Рис. 1. Структурные особенности иерархических данных, ведущие к неправильным выводам при применении МНК-регрессии – 1а: Разнонаправленность связи между переменными внутри групп и между группами. 1б: Арефакт «отсутствия эффекта» как результат разнонаправленности связей между группами. 1в: Обнаружение «эффекта» вследствие неучтенной группировки данных и коррелирующих ошибок. Рисунок авторов; панель 1а основана на [17, р. 15]

Структура многоуровневой модели

Простая иерархическая (многоуровневая) модель с одним предиктором на индивидуальном уровне и одним предиктором на групповом уровне выражается следующими образом.

$$\text{Уровень 1: } Y_{ij} = \beta_{0i} + \beta_{1j}X_{ij} + \varepsilon_{ij} \quad (1.1)$$

$$\text{Уровень 2: } \beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + \eta_{0j} \quad (1.2)$$

$$\beta_{1j} = \gamma_{11} + \gamma_{12}Z_j + \eta_{1j}, \quad \eta_{0j} \sim N(0, \gamma_{00}), \quad \varepsilon_{ij} \sim N(0, \sigma^2), \quad (1.3)$$

где субиндекс i относится к наблюдению, aj – к группе, соответственно, Y_{ij} – значение зависимой переменной для наблюдения i в стране j .

Уравнение для первого уровня (1.1) схоже с простой линейной моделью, однако субиндекс j показывает, что для каждой единицы второго уровня (страны, класса и т.д.) оценивается своя модель первого уровня.

Рассмотрим содержательную модель на примере доверия. В теориях доверия обычно считается, что генерализованное доверие – доверие незнакомцам – положительно связано с социальным капиталом, который определяется через высокий социальный класс и образование индивида [25]. Тем не менее в среднем прямая связь социального класса и образования с генерализованным доверием индивидов довольно слабая, так как баланс возможных рисков и вознаграждений доверия незнакомцам в разных странах различается. Применив многоуровневые модели, Т. Хамамура выявил, что социальный класс индивида положительно связан с генерализованным доверием только в богатых странах [26], причем для высшего класса в богатых странах связь существенно сильнее. Схожим образом, Н. Чаррон и Б. Ротстейн обнаружили, что связь образования и генерализованного доверия сильнее в странах с отсутствием фаворитизма в работе общественных институтов [27, р. 68], а М. Фредериксен с соавторами показали, что доверие положительно связано с генерализованным доверием только в

странах с низким уровнем коррупции [28, р. 305]. Эти результаты помогают объяснить, почему в Швеции среди самых образованных людей доверяют незнакомцам 73%, а в Перу – 8% [24]. Для межстрановых данных это означает, что у каждой страны в исследовании может быть свой уровень генерализованного доверия (β_{0j}) и свой случайный эффект образования (наклон β_{1j}), а ε_{ij} – ошибка модели для респондента i в стране j . Таким образом, мы моделируем дисперсию свободных членов уравнения для каждой страны, в литературе это сокращенно называется «моделью свободных интерсептов» («интерсепт» – точка пересечения регрессионной прямой с осью ординат).

Уравнение для второго уровня (1.2) показывает, что интерсепты и наклоны предикторов первого уровня являются функцией предикторов уровня страны и их вариации [16, р. 10]: β_{0j} – это интерсепт индивидуального уровня в стране j , γ_{00} – среднее значение генерализованного доверия с учетом предиктора второго уровня (коррупции) Z_j ; γ_{01} – эффект коррупции в стране Z_j , а η_{0j} – ошибка для страны j (если рассматривается вариация эффекта второго уровня, для каждой страны можно построить свой наклон эффекта коррупции).

Во втором уравнении странового уровня (1.3) моделируется эффект странового фактора на наклон эффекта индивидуального уровня (например, предполагается разный эффект образования в зависимости от уровня коррупции в стране) X_{ij} , где β_{1j} – эффект индивидуального предиктора в стране j , γ_{10} – среднее значение индивидуального эффекта с учетом странового фактора Z_j , γ_{11} – эффект странового фактора, а η_{1j} – ошибка по стране j . Двухуровневая модель целиком может описываться как такой системой уравнений, так и одним уравнением со смешанными (индивидуальными и страновыми) эффектами.

Коэффициенты регрессии в многоуровневых моделях бывают двух типов: фиксированные (fixed effects)¹ и рандомизированные,

¹ В модели с фиксированным эффектом регрессионные коэффициенты моделируются как одинаковые для всех единиц верхнего уровня [29, р. 590], например,

или случайные (random effects)¹. Оценки фиксированных эффектов выражают отношения между переменными независимо от того, к какой группе относятся наблюдения, как в обычной линейной регрессии. Для каждой группы можно также оценить случайный эффект, отражающий, насколько отношения между переменными на индивидуальном уровне отличаются от оценки фиксированного эффекта [4]. Данные повторяющихся исследований тоже могут быть представлены как иерархии, где моменты времени «вложены» в одни и те же единицы наблюдения [30, р. 124–125; 31, р. 134]. В лонгитюдных межстрановых исследованиях различают три уровня: «респондент», «год-страна» и «страна». Так, используя трехуровневую модель, М. Фэйрбразер показал на данных 1981–2008 гг., что уровень доверия, каким бы он ни был изначально, снижается в тех странах, где растет неравенство [30, р. 135]. При помощи трехуровневой модели Я. Мивес, используя данные 2002–2010 гг., обнаружил, что в европейских странах с низким уровнем гендерного равенства в сфере трудовых отношений женщины доверяют незнакомцам меньше, чем мужчины, тогда как в странах с высоким гендерным равенством такого эффекта не наблюдается [32].

В более сложных случаях может образоваться несколько естественных иерархий, которые перекрещиваются («невложенные» данные с перекрестной классификацией) [17, р. 206]. Эта ситуация наблюдается в опросах — например, когда респонденты

как одинаковый эффект пола респондента на уровень доверия во всех странах мира. Эта модель может быть реализована с помощью фиктивных переменных для единиц второго уровня (стран), метода робастных кластеризованных остатков (см. далее), а также в многоуровневой модели с фиксированными эффектами.

¹ Моделью со случайным, или рандомизированным, эффектом в многоуровневом моделировании называют модель, в которой эффект предиктора моделируется как варьирующий между единицами более высокого уровня [29, р. 593] — например, допускается наличие или отсутствие эффекта пола на доверие и различная направленность этого эффекта в зависимости от страны наблюдения.

представляют разные страны, но одновременно и разные религии или когда опрос проводится в разных странах, но отражает и конкретный год в каждой стране, и уникальный год в целом, когда могли происходить события международной важности: извержение вулкана, экономический кризис. Многоуровневое моделирование позволяет, таким образом, анализировать и срезные данные крупных межстрановых опросов, и лонгитюдные данные, в том числе в одной стране (при достаточно долгих временных рядах). Критерии для выбора этого метода анализа данных, помимо исследовательского вопроса, изложены далее.

Целесообразность многоуровневого моделирования

Цель многоуровневого моделирования, как и других видов регрессии, – объяснение дисперсии одной переменной на основе дисперсии других переменных. В данном случае моделируется переменная на самом низком, первом уровне данных, тогда как в качестве предикторов могут выступать переменные на всех уровнях идентифицированной исследователем иерархии. Многоуровневая модель особенно уместна там, где цель исследования – объяснить, каким образом характеристики единиц более высокого уровня связаны с поведением единиц низкого уровня [33, р. 219]. Дж. Гилл и Э. Вомэк указывают, что многоуровневое моделирование является мощным и гибким расширением классической (МНК) и обобщенной линейной модели, так как оно учитывает внутреннее агрегирование данных, в результате чего обычная модель дополняется структурным компонентом [34, р. 4]. Э. Гельман и Дж. Хилл при этом отмечают, что многоуровневое моделирование выступает компромиссом между МНК-моделью, где все наблюдения анализируются как единый массив без учета их принадлежности к какой-либо группе (fully pooled model), и построением совокупности отдельных моделей (fully unpooled models) [2, р. 6], когда оценивается серия регрессий для каждого объекта второго уровня.

Количество уровней в моделях определяется дизайном исследования (как правило, используются двух- и трехуровневые модели). В качестве теоретического и эмпирического примеров здесь и далее будем рассматривать самую простую, двухуровневую модель.

М. Стинберген и Б. Джонс выделяют две группы причин применять многоуровневое регрессионное моделирование для анализа данных [33, р. 219]. С одной стороны, это содержательная мотивация, с другой – статистическая. Говоря о содержательных причинах, авторы подчеркивают, что, во-первых, многоуровневое моделирование позволяет получить модель, исчерпывающе описывающую данные на разных уровнях, тогда как МНК-регрессия приводит к потере информации. Во-вторых, многоуровневые модели позволяют определить, влияют ли факторы второго уровня на вариацию эффектов независимых переменных первого уровня. В-третьих, многоуровневый анализ позволяет понять, можно ли распространять полученные результаты на другие группы и периоды времени, что особенно актуально для сравнительных межстрановых исследований.

Со статистической точки зрения, применение МНК-регрессий к данным, имеющим многоуровневую структуру, приводит к уменьшению дисперсии коэффициентов модели. Получаемые оценки при этом хоть и не смещены, но, как указывалось выше, могут быть ложно значимыми [35; 36].

Количественным критерием целесообразности применения многоуровневого регрессионного моделирования служит коэффициент внутриклассовой корреляции (ICC, intraclass correlation coefficient). Он показывает долю дисперсии зависимой переменной, которая объясняется разницей между группами. В других случаях данный коэффициент может использоваться как критерий надежности двух способов оценки одного и того же показателя с близкими средними значениями (например, собственная и чужая оценка уровня доверия индивида по одной и той же шкале) или доли дисперсии, объясняемой принадлежностью индивидов к группе в рамках однофакторного дисперсионного анализа

(ANOVA). В многоуровневом моделировании ИСС применяется, чтобы показать, насколько схожи между собой наблюдения внутри групп (скажем, стран) по значению зависимой переменной. Расчет коэффициента производится по формуле (2), где $\sigma^2_{between_class}$ – дисперсия интерсептов (β_{0j}), а $\sigma^2_{within_class}$ – дисперсия ошибок первого уровня (ϵ_{ij}). Коэффициент рассчитывается для так называемой пустой модели (формулы (3.1–3.2)), т. е. модели без предикторов.

$$ICC = \frac{\sigma^2_{between_class}}{\sigma^2_{between_class} + \sigma^2_{within_class}} \quad (2)$$

$$Y_{ij} = \beta_{0j} + \epsilon_{ij} \quad (3.1)$$

$$\beta_{0j} = \gamma_{00} + \eta_j \quad (3.2)$$

Величина ИСС связана с точностью оцениваемых в модели параметров, а также с эффектом дизайна, который показывает, насколько занижены стандартные ошибки в иерархических данных при применении МНК-регрессии. Эффект дизайна приблизительно равен $1 + (\text{средний размер кластера} - 1) * ICC$, если выборки по группам равны, и более – если они не равны [37, р. 87]. Как отмечает Й. Хокс, с ростом объема групповых выборок (например, от 5 членов семьи до 25 учеников класса в школе) и уровня ИСС эффект дизайна резко возрастает, поэтому игнорирование многоуровневой структуры данных в такой ситуации приводит к неверным оценкам параметров [36, р. 148]. В социальных науках наиболее часто встречаются значения $ICC = [0,00; 0,40]$ [17, р. 188], $ICC = [0,05; 0,40]$ [38, р. 168], $ICC = [0,00; 0,50]$ [39, р. 456]. При этом ИСС может стремиться к минимуму не только по причине отсутствия схожести наблюдений в группах, но и вследствие широкого разброса средних значений зависимой переменной между группами. По этой причине Э. Гельман и Дж. Хилл считают обоснованным применять многоуровневое моделирование начиная от значения $ICC = 0,05$ [2, р. 449]; в исследованиях, использующих имитационные модели, за минимум принимают $ICC = 0,05 - 0,10$ [21, р. 9; 39, р. 456].

Помимо ИСС, формальным критерием применения многоуровневого моделирования служит количество единиц наблюдения на групповом уровне. По оценке Б. Мутéна, при низких значениях ИСС на выборке в менее чем 50 стран многоуровневая модель не даст надежных результатов [39, р. 458]. Й. Хокс предлагает практическое правило 50/20: не менее 50 единиц на уровне групп и не менее 20 единиц наблюдения в каждой группе при моделировании межуровневых взаимодействий. А если исследователей интересует вариация эффектов группового уровня, Хокс предлагает аналогично действующее правило 100/10 [36, р. 151–152]. Некоторые исследования с использованием имитационных моделей показывают, что в линейной многоуровневой модели без взаимодействия наименьшее допустимое количество объектов второго уровня – 25, а для логистических моделей и 30–35 стран не всегда гарантируют надежных результатов [22, р. 14].

Таким образом, при соответствующей исследовательской задаче применение многоуровневых моделей оправдано при $ИСС > 0,05$ и количестве наблюдений на групповом уровне не менее 50. При меньшем количестве стран результаты могут быть неустойчивыми (а при $ИСС \approx 0$ усложнение модели до многоуровневой может быть неоправданным). В современных межстрановых исследованиях обеспечить участие 50 и более стран возможно не всегда; например, Европейское исследование ценностей охватывает менее 30 стран за один раунд. В таких случаях можно обратиться к байесовским методам оценки параметров многоуровневой модели¹ либо скорректировать исследовательский вопрос и использовать альтернативные многоуровневому моделированию методы. В общем случае мы рекомендуем использовать многоуровневое моделирование, когда имеется не менее 30 стран для анализа [21].

¹ Которые заслуживают отдельного введения, реализованы в различных вариантах и здесь не рассматриваются.

Оценка многоуровневой модели

Рассмотрим процесс построения двухуровневой линейной модели с межуровневым взаимодействием. Как и при оценке методом наименьших квадратов, процесс моделирования многоуровневых данных представляет собой последовательное тестирование усложняющихся моделей. Однако в данном случае модели различаются не только количеством независимых переменных и наличием взаимодействия, но и тем, допускается ли вариация интерсептов по группам и оцениваются ли наклоны регрессионной прямой для каждой группы отдельно (случайные эффекты/*random effects*) или же используется единый наклон прямой для всех групп (фиксированный эффект/*fixed effect*).

Первый шаг в построении многоуровневой модели – уже упоминавшаяся «пустая» модель (*empty model*, *null model*) (рис. 2а). Она дает общую информацию о дисперсии переменной, сопоставимую с ANOVA, позволяет рассчитать ICC и служит точкой отсчета при оценке качества других моделей. Значения интерсептов здесь могут рассматриваться как средние значения зависимой переменной для каждой группы.

Пустую модель (формула (3.1–3.2)) можно усложнить включением предикторов первого уровня (формулы (4.1–4.2)). Интерсепт β_0 в такой модели может изменяться по группам, в то время как наклон β_1 зафиксирован, т.е. эффект предиктора одинаков для всех групп (рис. 2б). Такая модель называется моделью с рандомизированным/случайным интерсептом (*random intercept model*). С содержательной точки зрения ее применение сопоставимо со сравнением групповых средних при контроле по индивидуальным параметрам респондентов в группах (*means-as-outcomes model* [1, p. 24–25]), например, моделирование среднего уровня генерализованного доверия с учетом социально-демографических признаков респондентов.

$$\text{Уровень 1: } Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \dots + \varepsilon_{ij} \quad (41)$$

$$\text{Уровень 2: } \beta_{0j} = \gamma_{00} + \eta_j \quad (42)$$

Альтернативным шагом усложнения «пустой» модели выступает модель со случайным эффектом β_{1j} (формулы (5.1–5.2)). Она предполагает общий для всех групп интерсепт, что является сильным допущением и редко используется в социальных науках. Такая модель технически вообразима при квазиэкспериментальном дизайне, когда специально подбираются случаи с одинаковым начальным значением параметра и моделируется различие эффекта [2, р. 283], или когда эмпирически было установлено отсутствие различий между интерсептами – если предположить, что практически вся дисперсия зависимой переменной объясняется моделью и необъясненной дисперсии почти не осталось (см. например: [40, р. 1207]):

$$\text{Уровень 1: } Y_{ij} = \beta_0 + \beta_{1j}X_{ij} + \dots + \varepsilon_{ij} \quad (51)$$

$$\text{Уровень 2: } \beta_{1j} = \gamma_{00} + \eta_j \quad (52)$$

Третьим шагом становится построение модели со случайным интерсептом β_{0j} и случайным эффектом β_{1j} , где каждая группа может иметь свой интерсепт и свой наклон регрессии с учетом вариации независимых переменных на первом и втором уровне (формулы (6.1–6.3), *рис.* 2г). Например, и средний уровень генерализованного доверия, и связь образования с доверием, и связь ВВП с доверием могут отличаться по странам.

$$\text{Уровень 1: } Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \dots + \varepsilon_{ij} \quad (6.1)$$

$$\text{Уровень 2: } \beta_{0j} = \gamma_{00} + \eta_{0j} \quad (6.2)$$

$$\beta_{1j} = \gamma_{10} + \eta_{1j} \quad (6.3)$$

На четвертом этапе в модель со случайными интерсептами и/или коэффициентами могут быть включены предикторы второго уровня (Z_j). Иными словами, в таких моделях появляется элемент объяснения межгрупповых различий в значении зависимой переменной (формулы (7.1–7.3), *рис.* 2д). Например, связь образования и доверия может быть разной в странах с высоким и низким доходом.

$$\text{Уровень 1: } Y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + \varepsilon_{ij} \quad (71)$$

$$\text{Уровень 2: } \beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + \eta_{0j} \quad (72)$$

$$\beta_{1j} = \gamma_{10} + \eta_{1j} \quad (73)$$

Пятый шаг – добавление межуровневого эффекта взаимодействия (формулы (1.1–1.3), *рис. 2е*). С его помощью можно проверить, влияет ли значение признака на групповом уровне на зависимую переменную на индивидуальном уровне, а также на связь между независимыми и зависимой переменными на индивидуальном уровне. Иными словами, речь о межуровневом (кросс-уровневом) взаимодействии (см. например: [31]). Возвращаясь к примеру с доверием, в данной модели можно оценить, есть ли последовательный паттерн связи образования и доверия в странах с высоким (А) и низким (Б) уровнем дохода. Модель с межуровневым взаимодействием, объясняющая вариацию индивидуальных эффектов по группам, называется полной [41, р. 151], или обобщенной [33, р. 224], многоуровневой моделью.

По мере наращивания вложенных моделей от простых к сложным возникает вопрос: какие предикторы следует моделировать со случайным эффектом, а какие – с фиксированным. Хотя этот выбор влияет на обобщаемость полученных результатов, до сих пор не существует общепринятой практики того, как принимать решение в этом вопросе [43, р. 272], поэтому авторы предлагают различные правила и стратегии.

Если анализ разведывательный или его цель – проверка фиксированного эффекта в данных, имеющих иерархическую структуру, можно оценивать эффекты, ориентируясь на сами данные (data-driven) [17, р. 102; 43, р. 274]. В таком случае можно получить надежные фиксированные коэффициенты, однако если в генеральной совокупности предполагаются случайные эффекты, модель с фиксированными эффектами всегда будет компромиссом между сходимостью модели и получением несмещенных оценок, так как исключение из модели наблюдаемых в генеральной совокупности случайных эффектов может привести к неверным выводам [44, р. 80].

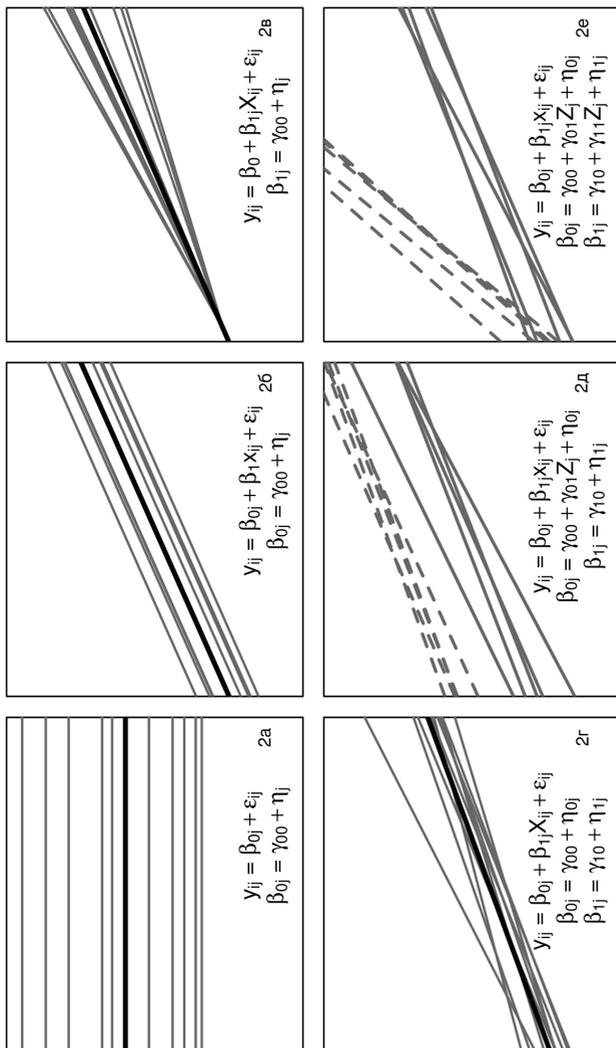


Рис. 2. Этапы многоуровневого моделирования. 2а: «пустая» модель (шаг 1), 2б: модель со случайным интерсептом (шаг 2), 2в: модель со случайным эффектом (альтернативный шаг 2), 2г: модель с случайными интерсептами и эффектами (шаг 3), 2д: модель с включением независимой переменной второго уровня (шаг 4), 2е: многоуровневая модель с межуровневым эффектом взаимодействия (шаг 5). Рисунок авторов; частично этапы отражены в [42, р. 727]

Если цель анализа – проверка теории, то предпочтителен выбор случайных эффектов на основе теории (model-driven). При этом некоторые авторы считают, что более надежно использовать «максимальную модель» [2, р. 246; 43, р. 274; 45, р. 35], где все коэффициенты на всех уровнях оцениваются как случайные на том основании, что, если эффект присутствует, он может проявляться между единицами наблюдения по-разному на разных уровнях. Таким образом, первое практическое правило многоуровневого моделирования – использовать случайные эффекты везде, где это возможно. Поскольку между странами почти всегда есть ненаблюдаемые устойчивые различия, игнорирование случайных эффектов на страновом уровне приведет к заниженным стандартным ошибкам и неоправданно оптимистичным выводам о наличии эффекта там, где его может не быть [45, р. 34].

Проблема такого подхода, особенно в случае сложной модели, возможная несходимость модели (non-convergence), когда не могут быть получены устойчивые параметры. Улучшает сходимость модели фиксирование ранее рандомизированных эффектов, однако это вряд ли приемлемое решение в случае проверки теории, так как оно ведет к смещенным оценкам дисперсии, особенно при гетероскедастичности остатков [44, р. 79]. Необоснованное фиксирование коэффициента может привести к неверной оценке стандартных ошибок и в межуровневом взаимодействии. По мнению Т. Снайдерса и Р. Боскера, отсюда следует общее правило: если исследователя интересует фиксированный эффект межуровневого взаимодействия $X*Z$, тогда X должен иметь свободный коэффициент на уровне Z (если его вариация значима) [17, р. 155]¹.

¹ Если в полной модели с межуровневой интеракцией было обнаружено, что вариация индивидуального эффекта объясняется группообразующей переменной, это может послужить основанием для фиксирования индивидуального эффекта (см. рис. 2е).

Если модель не сходится, можно также фиксировать эффекты контрольных переменных, второстепенных с точки зрения теории. Однако если эти переменные включены во взаимодействие с основными переменными, рекомендуется сначала проверить значимость рандомизированных эффектов контрольных переменных [43, р. 275]. Учитывая возможную несходимость, можно еще на этапе построения модели наращивать модель от самой простой, добавляя в нее один за другим все значимые случайные эффекты новых переменных.

Оценка качества многоуровневой модели

Для оценки качества многоуровневой модели невозможно рассчитать процент объясненной дисперсии (R^2), поскольку у предикторов разных уровней остатки распределяются неодинаково; то же касается и статистической значимости эффектов, поскольку на разных уровнях наблюдается разное количество степеней свободы. В связи с этим некоторые авторы отмечают нецелесообразность расчетов объясненной дисперсии в многоуровневых моделях [18, р. 119]. Другие авторы предлагают аналоги R^2 , называемые «псевдо- R^2 » и рассчитываемые для каждого уровня предикторов [17, р. 112; 46, р. 244] или для модели в целом [47]. Из них чаще других используется поуровневый псевдо- R^2 Т. Снайдерса и Р. Боскера (формулы (8.1–8.2), где σ^2 – дисперсия остатков на первом уровне для модели с предикторами, τ_{00} – дисперсия остатков на групповом уровне для модели с предикторами, σ^2_A – дисперсия остатков на индивидуальном уровне для «пустой» модели, τ_{00A} – дисперсия остатков на групповом уровне для «пустой» модели, n_j – среднее гармоническое количество наблюдений по группам) [17, р. 112–113]. Особенность этого подхода заключается в том, что получаемые значения псевдо- R^2 при добавлении предикторов могут быть отрицательными, что противоречит интуитивной интерпретации объясненной дисперсии (см. [17, р. 110]).

$$\text{Уровень 1: } R_1^2 = 1 - \frac{\hat{\sigma}^2 + \hat{\tau}_{00}}{\hat{\sigma}_A^2 + \hat{\tau}_{00A}} \quad (81)$$

$$\text{Уровень 2: } R_2^2 = 1 - \frac{\hat{\sigma}^2 / n_j + \hat{\tau}_{00}}{\hat{\sigma}_A^2 / n_j + \hat{\tau}_{00A}} \quad (82)$$

Псевдо- R^2 используется не всегда. Р. Сю предложил меру Ω^2 (омега-квадрат), рассчитываемый как: $1 - (\sigma^2/\sigma_A^2)$ [47, р. 3530]. Позже этот подход был раскритикован [48, р. 1905], хотя он и внедрен в некоторые статистические пакеты. Таким образом, в многоуровневой модели долю объясненной дисперсии можно считать лишь приблизительно и, чтобы ее верно интерпретировать, необходимо обращать внимание на то, какое именно определение используется в конкретном случае.

Ввиду сложности оценки объясненной дисперсии, для сравнения качества многоуровневых моделей часто используют показатели на основе функций максимального правдоподобия, которые используются для статистической оценки эффектов в многоуровневой регрессии. Снижение значений этих функций, по сравнению с «нулевой» моделью, указывает на улучшение объяснительной способности модели. Чаще всего используются функции полного максимального правдоподобия (FML) или ограниченного максимального правдоподобия (RML). Во многих случаях различия между двумя методами обычно малы [49, р. 41]. На основе FML используются такие показатели, как информационный критерий Акаике (AIC) и байесовский информационный критерий Шварца (BIC), подходящие для сравнения любых моделей с одинаковым количеством наблюдений: лучшей считается та модель, где значение критериев ниже [49, р. 51]. Для сравнения «вложенных» моделей (где конкретную модель можно получить из более общей путем фиксирования эффектов или изъятия предикторов) оценивается значимость хи-квадрата на основе теста отношения правдоподобия [49, р. 47–50].

Проверка надежности результатов

Для подтверждения надежности выводов необходима диагностика результатов моделирования. Такие классические методы диагностики, как анализ остатков, проверка на гомоскедастичность и выявление выбросов, применимы для многоуровневых регрессий, однако при наличии нескольких уровней анализа диагностика может производиться на каждом из них по отдельности.

Наглядным способом выявления влиятельных случаев второго уровня в двухуровневых регрессионных моделях может послужить построение диаграмм рассеяния для переменных, агрегированных на втором уровне [50, р. 174]. Поскольку экстремальные значения конкретных стран теоретически могут повлиять на итоговые коэффициенты, для проверки устойчивости модели также проводится диагностика влияния (расстояния Кука) и выбросов (DFBETAs) (обзор диагностик многоуровневых моделей см.: [51, р. 294–296]). Расстояния Кука дают один коэффициент влияния на модель в целом, а DFBETAs демонстрируют эффект удаления наблюдения второго уровня на каждый из коэффициентов.

Для расстояний Кука следует обратить внимание на коэффициенты, превышающие $4/n$ [50, р. 175; 52, р. 40–41], для диагностики выбросов DFBETAs – превышающие $2/\sqrt{n}$ по модулю, где n – количество наблюдений на данном уровне [50, р. 175]. При обнаружении влиятельных (highleverage) наблюдений, существенно изменяющих оценки параметров полной модели, рекомендуют дополнить модель межуровневого взаимодействия, удалить выбросы или использовать робастные способы оценки (бутстреппинг и др.) [51, р. 296]. При удалении влиятельных стран из итоговой модели следует показывать модели с ними и без них и по возможности интерпретировать различия в эффектах.

Многоуровневое моделирование в пакетах анализа данных

На сегодняшний день метод многоуровневого моделирования так или иначе реализован в популярных программах анализа данных, для каждой из которых существуют обучающие курсы (в том числе бесплатный обучающий модуль Европейского социального исследования ESS EduNet).

Специализированными программами для многоуровневого моделирования являются платный пакет HLM, широко используемый благодаря возможностям оценки моделей с порядковыми и бинарными зависимыми переменными и выбору алгоритма правдоподобия, а также пакет MLwiN с удобной реализацией байесовских оценок параметров. Полноценный расчет многоуровневых моделей возможен также в платных программах Mplus, SAS, Stata. В пакете IBM SPSS Statistics многоуровневое моделирование реализовано ограниченно: для оценки дисперсии используется исключительно Вальд-тест (в результате параметры часто завышены); также в нем невозможно моделировать бинарные, дискретные или порядковые переменные и др. [53, р. 386]. В современных учебниках по многоуровневому моделированию дается обзор этих программ (см.: [49, р. 39]). С нашей точки зрения, предпочтительна для ознакомления с методом именно *R* как свободно распространяемая программа с широкими возможностями выбора алгоритмов расчета и визуализации результатов (*R*-код для репликации представленного далее анализа см. в Приложении 1).

Эмпирический пример: влият ли ВВП на связь образования и доверия?

В качестве иллюстрации описанных выше принципов многоуровневого моделирования рассмотрим модель, предсказывающую уровень генерализованного доверия населения стран мира.

В межстрановых исследованиях доверия в качестве метода анализа достаточно часто используют многоуровневое моделирование (см.: [54; 55; 56]). Это связано с тем, что уровень доверия в обществе обусловлен не только свойствами доверяющего (его социальным положением, личным опытом, психологическими характеристиками и т. п.) [25, р. 95; 57, р. 33], но и средой, в которой доверие устанавливается [25, р. 97; 58]. При этом эффекты этих факторов в разных странах могут различаться.

Рассмотрим связь уровня образования и обобщенного доверия. Согласно теории Т. Ямагиши, доверие есть форма социального интеллекта, который можно определить как «способность распознавать и обрабатывать сигналы о рискованных социальных интеракциях» [59, р. 126], а образование является одним из институционализированных способов получения такого интеллекта. В литературе описаны несколько механизмов положительной связи образования с уровнем доверия. Так, ожидается, что с ростом образования индивиды приобретают навыки работы с информацией, и это ведет к повышению уровня их доверия [60]. В то же время в процессе получения образования индивид также становится более толерантным по отношению к людям из других социальных групп [61]. Несмотря на согласие ряда авторов по поводу положительной связи доверия и уровня образования респондента, эмпирические исследования показывают неоднозначные результаты [27, р. 60].

В представленной ниже многоуровневой модели объединены индивидуальные и страновые детерминанты доверия. Проверяется предположение о том, что взаимосвязь между доверием и уровнем образования различается в зависимости от экономического контекста.

Эмпирической базой данной модели выступает 6-я волна Всемирного исследования ценностей, проведенная в 2010–2014 гг. (WorldValuesSurvey, WVS) [24] и содержащая мнения о самых различных аттитюдах и ценностях. В анализе используются данные о 56 странах. Такое количество объектов второго уровня приемлемо для многоуровневого моделирования. Всего в базе

WVS представлены ответы 90350 человек; после удаления наблюдений, содержащих пропущенные значения в хотя бы одной из используемых в анализе переменных, наблюдений первого уровня осталось 80421¹.

Зависимая переменная – индекс генерализованного доверия, рассчитанный как сумма уровня доверия людям, с которыми респондент знаком лично, и уровня доверия людям, с которыми респондент встретился впервые. Получившийся индекс принимает значения от 2 до 8 (приложение 2). В качестве предикторов индивидуального уровня используются социально-демографические переменные: пол, возраст, децильная группа по доходу и уровню образования. Дескриптивные статистики представлены в *табл. 1*.

В качестве предиктора второго уровня используется внутренний валовой продукт (ВВП) на душу населения по паритету покупательной способности (данные Всемирного банка за год, соответствующий году сбора данных).

Процесс многоуровневого моделирования был реализован в несколько этапов. Сначала для проверки целесообразности многоуровневого анализа была построена «нулевая модель», которая показала, что ICC равен 0,115 (*табл. 2*, M0) и, таким образом, многоуровневый анализ оправдан. Интерсепты, соответствующие каждой из стран, в нулевой модели показывают средний уровень индекса генерализованного доверия по странам.

Далее в модель были добавлены фиксированные эффекты переменных первого уровня (*табл. 2*, M1–M5). Из *табл. 2* видно, что полученные эффекты индивидуальных переменных стабильны и значимы на уровне 0,001, а AIC и BIC минимальны в модели M5, включающей все предикторы. Все коэффициенты наклона в данных

¹ Целью данной статьи является показать основные шаги работы с методом, поэтому для простоты изложения диагностика пропущенных данных не рассматривается. Альтернативой удалению пропущенных значений в многоуровневой модели выступает множественное импутирование данных (подробнее см.: [62]).

Таблица 1
ОПИСАТЕЛЬНЫЕ СТАТИСТИКИ ПЕРЕМЕННЫХ МОДЕЛИ (WVS, 2010–2014)

Переменная	Доля, %	Среднее	Стандартное отклонение	Минимум	Максимум	Количество наблюдений
Индекс генерализованного доверия	–	4,88	1,33	2	8	80421
Доверие знакомым людям	–	2,94	0,81	1	4	80421
Доверие незнакомым людям	–	1,94	0,81	1	4	80421
Возраст	–	41,67	16,29	16	99	80421
Доход	–	4,88	2,11	1	10	80421
Пол: женский (фиктивная переменная)	52	–	–	0	1	41544
Пол: мужской (фиктивная переменная)	48	–	–	0	1	38877
Образование: начальное (фиктивная переменная)	23	–	–	0	1	18677
Образование: среднее (фиктивная переменная)	52	–	–	0	1	41722
Образование: высшее (фиктивная переменная)	25	–	–	0	1	20022

Примечание. Уровень общего дохода определялся по шкале 1–10, где 10 – самый высокий доход в стране; возраст измерялся как число полных лет респондента; шкала образования образована путем объединения категорий: «Никогда не учились», «Неполное начальное», «Начальное» (категория «начальное»); «Неполное среднее плюс профессиональные курсы (ПТУ, ФЗУ, РУ без среднего образования)», «Среднее специальное или профессионально-техническое образование (СПТУ, техникум, училище)», «Неполное общее среднее образование» (категория «среднее»); «Высшее» и «Неоконченное высшее образование (не менее трех курсов вуза)» (категория «высшее»).

моделях фиксированы, а интерсепт рандомизирован, т.е. учтено различие в среднем уровне доверия между странами, однако предполагается, что эффект индивидуальных предикторов на доверие везде одинаков. Такая модель позволяет сделать вывод о положительной связи возраста, дохода и образования с уровнем доверия; при этом мужчины доверяют другим людям в целом чаще, чем женщины.

В качестве следующего шага была проверена необходимость рандомизации эффекта образования (*табл. 3, М6*), в соответствии с гипотезой исследования.

Затем в модель был добавлен предиктор странового уровня (логарифм ВВП) (*табл. 3, М7*). На последнем шаге в модель был введен межуровневый эффект взаимодействия ВВП и образования респондента (*табл. 3, М8*). АIC и BIC принимают минимальное значение в модели с взаимодействием. Поскольку модели М5–М8 вложенные, для их сравнения можно использовать тест отношения правдоподобия, значения которого минимально в модели с взаимодействием. Псевдо- R^2 для М8 на первом уровне составил 2%, а на втором – 9%.

Многоуровневое моделирование предполагает оценку множества коэффициентов для каждой группы. Визуализация коэффициентов дает дополнительные возможности интерпретации результатов (*рис. 3*).

Определение значимости эффектов с помощью p -критерия в многоуровневых моделях является спорным¹, однако эффекты взаимодействия с доверительными интервалами можно визуализировать (*рис. 4*).

¹ В некоторых пакетах R для работы с многоуровневыми моделями значения стандартных ошибок вариации намеренно не указываются (`lmer4`), поскольку выборочное распределение часто асимметрично и стандартные ошибки не могут служить точной оценкой значимости эффекта; в других пакетах, применяемых для табличного представления результатов модели они конвенционально присутствуют (`sjp.lmer`, `stargazer`).

Таблица 2
ДЕТЕРМИНАНТЫ ГЕНЕРАЛИЗОВАННОГО ДОВЕРИЯ НА ИНДИВИДУАЛЬНОМ УРОВНЕ
(данные WVS)

	M0	M1	M2	M3	M4	M5
	Фиксированная часть модели					
Интерсепт	4,86*** (0,06)	4,86*** (0,06)	4,89*** (0,06)	4,66*** (0,06)	4,81*** (0,06)	4,63*** (0,06)
Возраст		0,06*** (0,00)				0,08*** (0,00)
Пол: женский			-0,06*** (0,00)			-0,05*** (0,00)
Доход				0,04*** (0,00)		0,04*** (0,00)
Образование: среднее					0,03* (0,01)	0,05*** (0,01)
Образование: высшее					0,15*** (0,01)	0,15*** (0,01)

Примечания. Количество наблюдений: респондентов – 80421, стран – 56, в скобках указано значение стандартной ошибки; * $p < 0,05$, *** $p < 0,001$.

Окончание табл. 2

	M0	M1	M2	M3	M4	M5
	Рандомизированная часть модели					
Дисперсия интерцептов	0,204	0,199	0,204	0,200	0,200	0,191
σ^2	1,5685	1,565	1,568	1,562	1,565	1,555
$\tau_{\text{ур}}$, уровень страны	0,2039	0,199	0,204	0,200	0,200	0,191
Ω^2		0,109	0,108	0,112	0,109	0,116
AIC		264565	264565	264365	264563	264009
BIC		264602	264602	264402	264609	264083
псевдо- R^2 уровень 1		0,005	0,000	0,006	0,004	0,015
псевдо- R^2 уровень 2		0,024	-0,001	0,019	0,019	0,063

Таблица 3
ПОСТРОЕНИЕ МНОГОУРОВНЕВОЙ МОДЕЛИ ГЕНЕРАЛИЗОВАННОГО ДОВЕРИЯ (данные WVS)

	M5	M6	M7	M8
	Фиксированная часть модели			
Интерсепт	4,63*** (0,06)	4,62*** (0,06)	4,61*** (0,06)	4,61*** (0,06)
Возраст	0,08*** (0,00)	0,08*** (0,00)	0,08*** (0,00)	0,08*** (0,00)
Пол: женский	-0,05*** (0,01)	-0,05*** (0,01)	-0,05*** (0,01)	-0,05*** (0,01)
Доход	0,04*** (0,00)	0,04*** (0,00)	0,04*** (0,00)	0,04*** (0,00)
Образование: среднее	0,05*** (0,01)	0,05* (0,02)	0,05* (0,02)	0,05* (0,02)
Образование: высшее	0,15*** (0,01)	0,17*** (0,04)	0,17*** (0,04)	0,16*** (0,04)
log(ВВП)			0,17** (0,05)	0,09 (0,06)
Образование: среднее * log(ВВП)				0,50* (0,02)
Образование: высшее * log(ВВП)				0,10** (0,04)

Примечания. Количество наблюдений: респондентов – 80421, стран – 56, в скобках указано значение стандартной ошибки; * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$.

Окончание табл. 2

	M5	M6	M7	M8
Рандомизированная часть модели				
Дисперсия интерсептов	0,191	0,195	0,193	0,186
дисперсия коэффициента: среднее образование		0,020	0,020	0,018
дисперсия коэффициента: высшее образование		0,066	0,066	0,057
σ^2	1,555	1,546	1,547	1,547
$\tau_{\text{юр.}}$ уровень страны	0,191	0,195	0,193	0,186
Ω^2	0,116	0,121	0,121	0,121
AIC	264009	263751	263744	263741
BIC	264083	263872	263874	263890
-2LogLikelihood (χ^2 , вложенные модели)	263993	263725	263716	263709
псевдо- R^2 уровень 1	0,015	0,018	0,018	0,022
псевдо- R^2 уровень 2	0,063	0,044	0,053	0,087

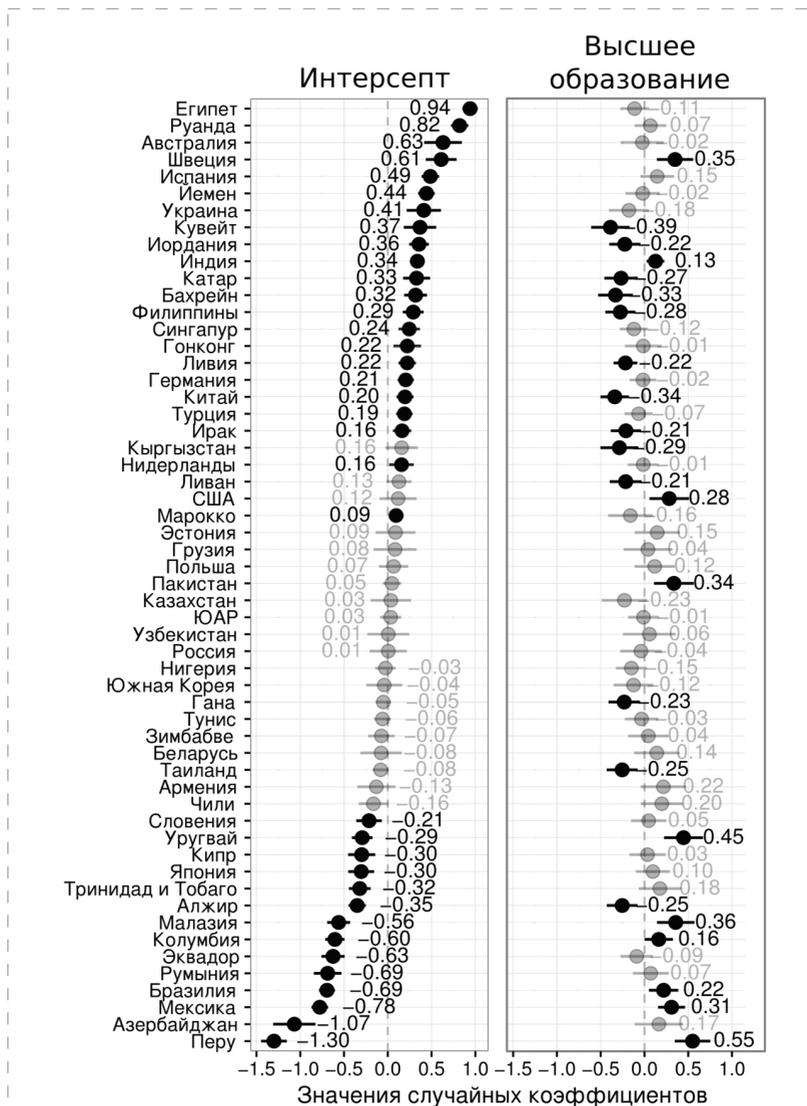


Рис. 3. Рандомизированные коэффициенты высшего образования в модели с межуровневым взаимодействием

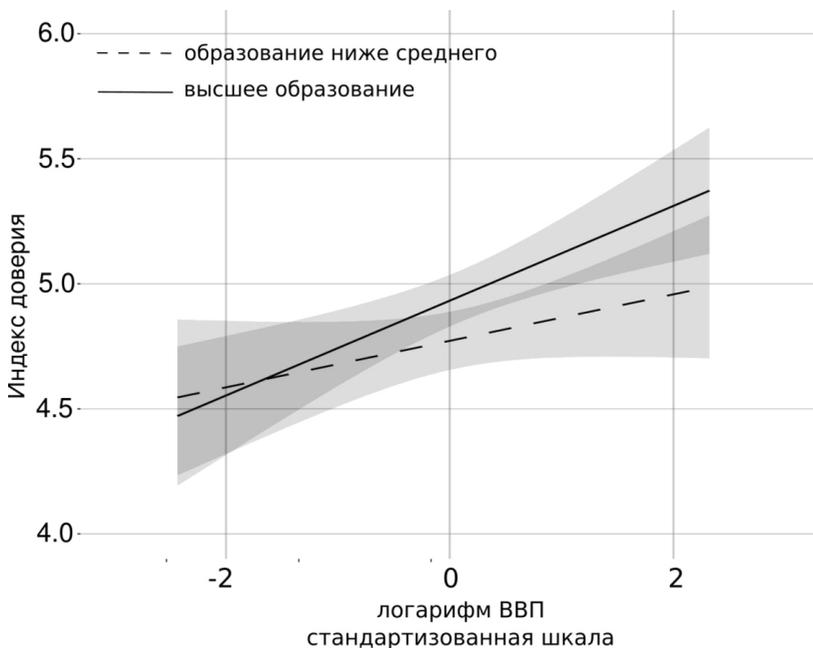


Рис. 4. Связь индекса доверия с уровнем образования и логарифмом ВВП с учетом эффекта их взаимодействия

Примечание: для наглядности показаны линии наклона только для начального и высшего образования.

На рис. 3 видно, что разница в уровне доверия между людьми с высшим образованием и образованием ниже среднего растет по мере увеличения ВВП страны. Построение доверительных интервалов показывает, что наблюдаемая связь образования и обобщенного доверия в зависимости от ВВП недостаточно велика, что отражает и псевдо- R^2 модели. Поскольку в данном случае модель с взаимодействием демонстрирует лучшие показатели по AIC, BIC и отношению правдоподобия, принимаем ее в качестве окончательной на данном этапе. Дальнейшая проверка теории Ямагиши может быть связана с использованием более специфичных предикторов, чем ВВП (см.: [27]).

Диагностика и устойчивость модели

Первый способ проверки устойчивости полученных результатов модели – построение диаграммы рассеяния, показывающей взаимосвязь значений логарифма ВВП и среднестранового индекса доверия (рис. 5).

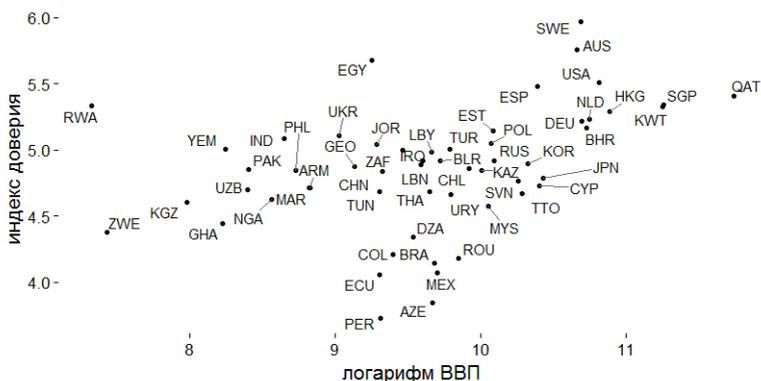


Рис. 5. Взаимосвязь индекса доверия и логарифма ВВП на межстрановом уровне

Примечание. Названия стран даны согласно кодам ISO3.

Из рис. 5 видно, что Руанда (RWA) является потенциально влиятельным случаем, так как при низком значении логарифма ВВП значение индекса доверия в стране достаточно велико. Поскольку сильно отклоняющиеся от линии регрессии значения могут повлиять на итоговые коэффициенты, рассчитываются статистики влияния и выбросов с помощью расстояний Кука и DFBETAs (обзор диагностик многоуровневых моделей см.: [51]).

По другой диагностике, расстояниям Кука, дающей общую оценку влияния наблюдения по всем предикторам, наиболее влиятельными в данном случае будут значения, превышающие $4/56 \approx 0,071$ (см. график в Приложении 3). Повторив модель по от-

дельности без каждой из стран со значением 0,071 и выше, видим, что исключение из модели Руанды делает значимым (положительный) эффект ВВП; исключение ЮАР (ZAF), Нигерии (NGA) или Ливии (LBY) существенно коэффициентов не изменяет. При расчете модели без США (USA) или Швеции (SWE) незначимым становится эффект среднего образования: таким образом, включение этих стран влияет на обнаружение положительной связи среднего образования с доверием в общей модели.

При диагностике модели с помощью другого обобщающего показателя, DFBETAs, пороговым служит значение, превышающее

$|2/\sqrt{56}| \approx |0,267|$. В таблице результатов¹ заметны особенности как уже упомянутых, так и других стран: например, положительные эффекты взаимодействия среднего и высшего образования в Швеции и США или отрицательные эффекты взаимодействия ВВП и образования в Бахрейне (BHR), Пакистане (PAK), Катаре (QAT). Подобные результаты детальнее показывают эффекты в разных странах и должны приниматься во внимание при общей оценке и описании модели.

В целом, полученные в итоговой модели результаты устойчивы для 55 стран при исключении «раскачивающего» модель ВВП Руанды; в то же время эффект взаимодействия среднего образования с ВВП на доверие меняется при исключении влиятельных наблюдений, а эффект взаимодействия высшего образования с ВВП, наоборот, сохраняется. Таким образом, независимо от уровня экономического развития наблюдается положительная связь высшего образования и обобщенного доверия, а изначальная гипотеза об определении положительной связи уровня образования и обобщенного доверия ВВП требует дальнейшего уточнения.

¹ Доступны по запросу у авторов.

Альтернативы многоуровневому моделированию

Несмотря на изложенные выше достоинства многоуровневого регрессионного моделирования, иерархические данные не всегда требуют применения данного метода. Существует ряд методов, позволяющих получать результаты, схожие с многоуровневым моделированием при исследовательских вопросах, не требующих оценки случайных межстрановых эффектов или при недостаточном количестве наблюдений второго уровня и других нарушениях предположений многоуровневого моделирования.

Наиболее масштабной альтернативой выступают методы оценки эффектов кластеризации без разделения вариации на уровни, из которых рассмотрим два популярных: метод робастных кластеризованных остатков (cluster robust standard errors)¹ и метод обобщенных оценочных уравнений (generalized estimating equations). Эти методы моделируют эффекты для «глобального среднего» наблюдения (population-averaged), тогда как многоуровневые модели позволяют проследить изменения на уровне конкретной страны (subject-specific).

Метод робастных кластеризованных остатков, доминирующий в эконометрике (обзор см.: [63]), используется при анализе панельных данных (например, о работе одних и тех же компаний во времени): поскольку в таком случае нарушается допущение о независимости остатков, для вычисления стандартных ошибок регрессионных эффектов делаются поправки на основе скорректированной матрицы вариаций-ковариаций. Данный метод, как и многоуровневое моделирование, допускает кластеризацию по двум основаниям и трехуровневые модели (называемые multiway clustering). По сравнению с многоуровневым моделированием,

¹ Также они называются панельными робастными/ состоятельными/ устойчивыми к кластеризации стандартизированными ошибками методом сэндвич-оценки.

метод кластеризованных остатков проще в реализации и не предполагает допущений о распределении случайных эффектов [4]. Сравнительные преимущества моделей с кластеризованными остатками: большая сходимость моделей, чем в многоуровневых моделях [64, р. 455], более интуитивная оценка качества модели (R^2) и интерпретация результатов, особенно в случае дискретной зависимой переменной [4]. Преимуществами многоуровневого моделирования перед этим методом являются оценка параметров для каждой страны с учетом страновых предикторов и большая устойчивость к разному количеству наблюдений по странам. В случае если гипотезы не связаны с межгрупповой вариацией эффектов или кластеризацию остатков необходимо учесть только для точной оценки индивидуальных параметров, или иерархии в данных возникают как побочное следствие процедуры сбора данных, рекомендуется использовать метод робастных кластеризованных остатков [4].

Другим альтернативным методом этой группы, применяемым в биостатистике, эпидемиологии и некоторых видах когортных исследований, служат обобщенные оценочные уравнения (generalized estimating equations). Метод основан на повторяющемся анализе остатков для оценки параметров и применении робастных кластеризованных стандартных остатков на последнем этапе [4]. По сравнению с методом робастных остатков, обобщенные оценочные уравнения не допускают перекрестной кластеризации и более сложны в оценке качества модели (используется QIC, информационный критерий квазиправдоподобия), однако их результат более устойчив при разных размерах кластеров, высокой внутриклассовой корреляции (в лонгитюдах) и большом (>50) количестве кластеров. По сравнению с многоуровневыми моделями, обобщенные оценочные уравнения не позволяют моделировать случайные эффекты, однако они быстрее и проще в оценке как непрерывных, так и дискретных зависимых переменных [4]. В целом, два описанных альтернативных метода быстрее решают

проблему кластеризованных остатков, когда те скорее мешают анализу, чем представляют содержательный интерес; эти методы также более устойчивы при небольших (менее 5 наблюдений) размерах групп, что отчасти может объяснить их популярность в эпидемиологии и менеджменте. В то же время «если выборки обширны и кластеры значительно различаются по количеству наблюдений, многоуровневые модели могут быть лучшим способом работы с данными» [4] – что характерно для межстрановых опросов.

Внутри парадигмы иерархического анализа («раскладывания» вариации по уровням данных) также существует несколько альтернатив многоуровневому моделированию.

Во-первых, учет групповых различий возможен при включении в модель фиктивной переменной, указывающей, к какому из объектов второго уровня относится каждое наблюдение. В результате исследователь получает ряд коэффициентов (их число равно $j - 1$, где j – число объектов второго уровня)¹, показывающих, насколько отличается значение интереспта в каждой группе от значения интереспта для опорной категории. Статистическая надежность результатов делает этот подход популярным, однако он не позволяет получить богатые с теоретической точки зрения выводы [33, р. 220], которые можно было бы обобщить на все объекты второго уровня, но может быть полезным для несмещенной оценки модели в конкретной стране и анализе ее отличий от других стран.

Во-вторых, альтернативой многоуровневому моделированию является включение в МНК-модель предиктора второго уровня. В отличие от включения фиктивных переменных, данный подход позволяет моделировать эффекты второго уровня, но не учитывает отличий конкретных стран (например, если две страны имеют одинаковый уровень ВВП, в такой модели эти страны будут неразличимы). При этом данный подход позволяет учесть разницу интересптов в зависимости от значения страновой переменной.

¹ В рассмотренной полной модели их было бы $56 - 1 = 55$.

Такие модели могут дать содержательно богатые результаты, однако, по сравнению с многоуровневым моделированием, они менее надежны статистически, так как игнорируют реальное (небольшое) число степеней свободы предиктора на уровне страны [33, p. 220–221]. Еще одним способом анализа многоуровневых данных внутри этой парадигмы служит агрегирование всех показателей до уровня стран. В таком случае исследователь получает общее количество наблюдений, равное количеству стран, и полностью теряет информацию об индивидах. Такой подход, хоть и дает статистически надежные результаты при достаточном количестве наблюдений, может привести к неверным интерпретациям, связанным с экологической ошибкой, когда эффекты одного и того же показателя на индивидуальном уровне и в агрегированном виде на групповом уровне отличаются (см.: [17, p. 83-84; 65, p. 434]).

В целом, многоуровневое моделирование оказалось теоретически наиболее уместным методом анализа в том случае, когда многоуровневая структура данных представляет исследовательский интерес, а не является нарушением допущений анализа, требующим корректировки. Каждая из рассмотренных альтернатив многоуровневому моделированию должна применяться сообразно цели исследования.

Заключение

Обзор распространенных методов анализа кластеризованных данных показывает, что в естественных и социальных науках зачастую используются разные методы для решения схожих задач [4]. Целью данной работы было объяснение целей и способа проведения многоуровневого регрессионного моделирования на основе межстрановых опросных данных. На примере генерализованного доверия были представлены процедуры построения и оценки двухуровневой линейной регрессионной модели для данных с иерархической структурой, где индивиды и страны выступают как различные уровни кластеризации информации.

Были продемонстрированы необходимые условия применения многоуровневых моделей и их преимущества по сравнению с регрессионным моделированием методом наименьших квадратов, использованием фиктивных переменных, агрегированием индивидуальных предикторов, применением робастных кластеризованных остатков и обобщенных оценочных уравнений. На примере двухуровневой модели рассмотрены различные спецификации многоуровневых моделей, описаны шаги моделирования и способы оценки качества итоговой модели.

В эмпирической части работы на межстрановых сравнительных данных WVS показано, каким образом можно учитывать эффекты взаимодействия страновых и индивидуальных переменных в моделировании связи образования и уровня генерализованного доверия в зависимости от ВВП страны. В Приложении 1 представлен код в открытом статистическом пакете *R* для репликации описанной модели, ее дальнейшего развития и использования.

Наконец, была представлена двухуровневая линейная модель с межуровневым взаимодействием без перекрестной классификации наблюдений. Дальнейшее усложнение и развитие многоуровневой модели, помимо моделирования категориальной переменной, возможно в следующих направлениях. Во-первых, моделирование бинарных, порядковых и счетных зависимых переменных. Во-вторых, увеличение количества уровней, что может быть обусловлено повторными исследованиями или включением уровня регионов между индивидом и страной. В-третьих, добавление перекрестных классификаций, связанное с наличием нескольких оснований группировки данных (например, религия, профессия, страна). В-четвертых, применение различных функций максимального правдоподобия, а также байесовской оценки с помощью марковских цепей Монте-Карло (MCMC) и альтернативного им метода интегрированных вложенных аппроксимаций Лапласа (INLA) [66] для получения доверительных интервалов полученных коэффициентов.

Многоуровневое моделирование позволяет успешно отвечать на исследовательские вопросы, связанные с эффектом на групповом уровне на связь между переменными на уровне индивида, а также в любых других случаях данных с иерархической структурой, будь то срезовой или лонгитюдный способ сбора данных. Статистически надежные выводы при многоуровневом моделировании межстрановых сравнительных данных могут быть получены при минимальном эмпирически приемлемом количестве наблюдений второго уровня (30 – для непрерывной зависимой переменной, больше – для категориальной) и уровне «похожести» респондентов в одной группе, по сравнению с различиями респондентов между группами, ИСС. Многоуровневое моделирование используется наряду с другими способами учета автокорреляции остатков, популярных в разных науках, однако обладает рядом преимуществ, которые могут компенсировать более жесткие предположения этого метода.

ЛИТЕРАТУРА

1. *Raudenbush S.W., Bryk A.S.* Hierarchical Linear Models: Applications and Data Analysis Methods. Thousand Oaks: Sage, 2002.
2. *Gelman A., Hill J.* Data Analysis Using Regression and Multilevel/Hierarchical Models. New York: Cambridge Univ. Press, 2006.
3. *Mills M.C., Präg P.* Methodological Advances in Cross-National Research: Multilevel Challenges and Solutions // *European Sociological Review*. 2016. N. 1. P. 1–2.
4. *McNeish D., Stapleton L.M., Silverman R.D.* On the Unnecessary Ubiquity of Hierarchical Linear Modeling // *Psychological Methods*. 2016. Advance online publication. URL: <http://dx.doi.org/10.1037/met0000078> (date of access: 30.11.2016).
5. *Руднев М.Г.* Влияние страны происхождения и страны проживания на базовые ценности (по результатам изучения европейских мигрантов) // XIII Апрельская международная научная конференция по проблемам развития экономики и общества: В 4 кн. Кн. 2. С. 591-604. URL: <https://publications.hse.ru/chapters/67471834> (дата обращения: 30.04.2017).
6. *Сальникова Д.В.* Моделирование взаимосвязи между субъективным экономическим благополучием граждан и поддержкой институтов социального государства в странах ЕС // *Прикладная эконометрика*. 2014. №1 (33) С. 71–89.

7. Русинова Н.Л., Сафронов В.В. Здоровье в странах Европы: статусные неравенства и влияние социального капитала // Социологический журнал. 2015. Т. 21. № 4. С. 34–60.

8. Костенко В.В. Гендерные установки мигрантов-мусульман в западной Европе: многоуровневый неиерархический анализ // Мониторинг общественного мнения: экономические и социальные перемены. 2017. № 1 (137). С. 97–120.

9. Рейзингер У.М., Пачеко Д. Массовые демократические ценности и демократия в российских регионах // Мониторинг общественного мнения: экономические и социальные перемены. 2014. № 5 (123). С. 12–30.

10. Седова Е.Н., Стебунова О.И., Ушатова С.Т. Эконометрическое моделирование стоимости жилой недвижимости на региональном уровне: иерархический подход // Интеллект. Инновации. Инвестиции. 2016. Т. 4. С. 51–56.

11. Титкова В.В., Ивановишна В.А., Александров Д.А. Популярность школьников и образовательная среда школы // Вопросы образования. 2013. № 4. С. 145–167.

12. Уваров А.Г., Ястребов Г.А. Социально-экономическое положение семей и школа как конкурирующие факторы образовательных возможностей: ситуация в России // Мир России. Социология. Этнология. 2014. Т. 23. № 2. С. 103–132.

13. Кузьмина Ю.В. Эффект «большой лягушки в маленьком пруду»: всегда ли хорошо, если ребенок обучается в сильном классе? // Психология. Журнал Высшей школы экономики. 2016. Т. 13. № 4. С. 712–740.

14. Четвериков А.А. Линейные модели со смешанными эффектами в когнитивных исследованиях // Российский журнал когнитивной науки. 2015. Т. 2. № 1. С. 41–51.

15. Алексеева С.В., Слюсарь Н.А. Эффект длины при парафовеальной обработке слов во время чтения // Вестник Томского государственного университета. Филология. 2017. № 45. С. 5–29.

16. Luke D.A. *Multilevel Modelling*. Thousand Oaks: Sage, 2004.

17. Snijders T.A.B., Bosker R.J. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel*. Thousand Oaks: Sage, 2012.

18. Kreft I.G.G., Leeuw J. de. *Introducing Multilevel Modelling*. London: Sage, 1998.

19. Spencer N.H. Models for Value-Added Investigations of Teaching Styles Data // *Journal of Data Science*. 2008. N. 1. P. 33–51.

20. Geiser C. *Data Analysis with Mplus*. New York: Guilford Press, 2013.

21. Bryan M.L., Jenkins S.P. Multilevel Modeling of Country Effects: A Cautionary Tale // *European Sociological Review*. 2016. N. 1. P. 3–22.

22. Subramanian S.V. et al. Revisiting Robinson: The Perils of Individualistic and Ecologic Fallacy // *International Journal of Epidemiology*. 2009. N. 2. P. 342–360.

23. Bartels B. Beyond ‘Fixed versus Random Effects’: A Framework for Improving Substantive and Statistical Analysis of Panel, Time-Series Cross-Sectional, and Multilevel Data // *Quantitative Research in Political Science* / Ed. R. Franzese, SAGE Publications, 2008.

24. *World Values Survey Association*. World Values Survey Wave 6 2010–2014 Official Aggregate. v.20150418 [Электронный ресурс]. URL: <http://www.worldvaluessurvey.org/WVSDocumentationWVL.jsp> (date of access: 30.09.2016)
25. *Delhey J., Newton K.* Who Trusts?: The Origins of Social Trust in Seven Societies // *European Societies*. 2003. N. 2. P. 93–137.
26. *Hamamura T.* Social Class Predicts Generalized Trust but Only in Wealthy Societies // *Journal of Cross-Cultural Psychology*. 2012. N. 3. P. 498–509.
27. *Charron N., Rothstein B.* Does Education Lead to Higher Generalized Trust? The Importance of Quality of Government // *International Journal of Educational Development*. 2016. N. 50. P. 59–73.
28. *Frederiksen M., Larsen C.A., Lolle H.L.* Education and trust Exploring the association across social relationships and nations // *Acta Sociologica*. 2016. Vol. 59. N. 4. P. 294–308.
29. *Diez Roux A.V.* A glossary for multilevel analysis. *Journal of epidemiology and community health*. 2002. Vol. 56. N. 8. P. 588–594.
30. *Fairbrother M.* Two Multilevel Modeling Techniques for Analyzing Comparative Longitudinal Survey Datasets // *Political Science Research and Methods*. 2014. Vol. 2. N. 1. P. 119–140.
31. *Bell A., Jones K.* Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods*. 2015. Vol. 3. N. 1. P. 133–153.
32. *Mewes J.* Gen(d)eralized Trust: Women, Work, and Trust in Strangers // *European Sociological Review*. 2014. N. 3. P. 373–386.
33. *Steenbergen M.R., Jones B.S.* Modeling Multilevel Data Structures // *American Journal of Political Science*. 2002. N. 1. P. 218–237.
34. *Gill J., Womack A.J.* The Multilevel Model Framework // *The SAGE Handbook of Multilevel Modeling* / Eds. M.A. Scott, J.S. Simonoff, B.D. Marx. Thousand Oaks: SAGE Publications, 2013.
35. *Tate R.L., Wongbunhit Y.* Random versus Nonrandom Coefficient Models for Multilevel Analysis // *Journal of Educational and Behavioral Statistics*. 1983. N. 2. P. 103–120.
36. *Hox J.J.* Multilevel Modeling: When and Why // *Classification, Data Analysis, and Data Highways* / Eds. I. Balderjahn, R. Mathar, M. Schader. Berlin, Heidelberg: Springer, 1998.
37. *Maas C.J., Hox J.J.* Sufficient Sample Sizes for Multilevel Modeling // *Methodology*. 2005. N. 3. P. 86–92.
38. *Preacher K.J., Zhang Z., Zychur M.J.* Alternative Methods for Assessing Mediation in Multilevel Data: The Advantages of Multilevel SEM // *Structural Equation Modeling*. 2011. N. 2. P. 161–182.
39. *Muthén B.* Latent Variable Modeling of Longitudinal and Multilevel Data // *Sociological Methodology*. 1997. N. 1. P. 453–480.

40. Erceg V. et al. An Empirically Based Path Loss Model for Wireless Channels in Suburban Environments // IEEE Journal on Selected Areas in Communications. 1999. N. 7. P. 1205–1211.

41. Austin P.C., Goel V., Walraven C. van. An Introduction to Multilevel Regression Models // Canadian Journal of Public Health. 2001. N. 2. P. 150–154.

42. Hofmann D.A. An overview of the logic and rationale of hierarchical linear models // Journal of management. 1997. Vol. 23. N. 6. P. 723–744.

43. Barr D.J. et al. Random Effects Structure for Confirmatory Hypothesis Testing: Keep It Maximal // Journal of Memory and Language. 2013. N. 3. P. 255–278.

44. Townsend Z. et al. The Choice Between Fixed and Random Effects // The SAGE Handbook of Multilevel Modeling / Eds. M.A. Scott, J.S. Simonoff, B.D. Marx. Thousand Oaks: SAGE Publications, 2013.

45. Schmidt-Catran A.W., Fairbrother M. The Random Effects in Multilevel Models: Getting Them Wrong and Getting Them Right // European Sociological Review. 2015. N. 1. P. 23–38.

46. Gelman A., Pardoe I. Bayesian Measures of Explained Variance and Pooling in Multilevel (Hierarchical) Models // Technometrics. 2006. N. 2. P. 241–251.

47. Xu R. Measuring Explained Variation in Linear Mixed Effects Models // Statistics in Medicine. 2003. N. 22. P. 3527–3541.

48. Orelie J.G., Edwards L.J. Fixed-effect variable selection in linear mixed models using R2 statistics // Computational Statistics & Data Analysis. 2008. N. 52. P. 1896–1907.

49. Hox J.J. Multilevel Analysis: Techniques and Applications. Hove: Routledge, 2010.

50. Meer T. van der, Grotenhuis M.T., Pelzer B. Influential Cases in Multilevel Modeling: A Methodological Comment // American Sociological Review. N. 1. P. 173–178.

51. Aguinis H., Gottfredson R.K., Joo H. Best-Practice Recommendations for Defining, Identifying, and Handling Outliers // Organizational Research Management. 2013. T. 16. N. 2. P. 270–301.

52. Nieuwenhuis R., Grotenhuis H.F. te, Pelzer B.J. influence.ME: Tools for Detecting Influential Data in Mixed Effects Models // The R-Journal. 2012. N. 2. P. 38–47.

53. Hayes A.F. A Primer on Multilevel Modeling // Human Communication Research. 2006. N. 4. P. 385–410.

54. Delhey J., Welzel C. Generalizing Trust: What Extends Trust from Ingroups to Outgroups? // World Values Research. 2012. N. 3. P. 45–69.

55. Gheorghiu M.A., Vignoles V.L., Smith P.B. Beyond the United States and Japan: Testing Yamagishi's Emancipation Theory of Trust across 31 nations // Social Psychology Quarterly. 2009. N. 4. P. 365–383.

56. Robbins B.G. Neither Government nor Community Alone: A Test of State-Centered Models of Generalized Trust // Rationality and Society. 2011. N. 3. P. 304–346.

57. *Алмакаева А.М.* Измерение генерализованного (обобщенного) доверия в кросскультурных исследованиях // Социологические исследования. 2014. № 11. С. 32–43.
58. *Alesina A., La Ferrara E.* Who Trusts Others? // Journal of Public Economics. 2002. No. 2. P. 207–234.
59. *Yamagishi T.* Trust as a Form of Social Intelligence // Trust in Society / Ed. K.S. Cook. New York: Russell Sage Foundation, 2001.
60. *Keefer P., Knack S.* Social Capital, Social Norms and the New Institutional Economics // Handbook of New Institutional Economics / Ed. C. Ménard, M.M. Shirley. Berlin, Heidelberg: Springer, 2008.
61. *Helliwell J.F., Putnam R.D.* Education and Social Capital // Eastern Economic Journal. 2007. N. 1. P. 1–19.
62. *Фабрикант М.С.* Модель-ориентированный подход к отсутствующим значениям: множественная импутация в многоуровневой регрессии посредством R (на примере анализа опросных данных) // Социология: методология, методы, математическое моделирование. 2015. № 41. С. 7–29.
63. *Petersen M.A.* Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches // Review of Financial Studies. 2009. Vol. 22. P. 435–480.
64. *Primo D.M., Jacobsmeier M.L., Milyo J.* Estimating the impact of state policies and institutions with mixed-level data // State Politics & Policy Quarterly. 2007. Vol. 7. N. 4. P. 446–459.
65. *Gelman A.* Multilevel (hierarchical) modeling: what it can and cannot do // Technometrics. 2006. Vol. 48. N. 3. P. 432–435.
66. *Rue H., Martino S., Chopin N.* Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations // Journal of the royal statistical society: Series b (statistical methodology). 2009. Vol. 71. N. 2. P. 319–392.

Приложение 1

КОД ДЛЯ РЕПЛИКАЦИИ МОДЕЛЕЙ¹

```
library(foreign)
library(lme4)
library(sjPlot)
library(ggplot2)
library(ggrepel)
library(influence.ME)
data<- read.spss(«WV6_Data_spss_v_2016_01_01.sav», to.data.frame =
TRUE)
datac<- read.csv(«clevel.csv», header = TRUE)
vars<- c(
  «V2», #код страны
  «V104», #Вопрос «How much you trust: People you know personally»
  «V105», #Вопрос «How much you trust: People you meet for the first
time»
  «V239», #шкала доходов
  «V240», #пол
  «V242», #возраст
  «V253», #размер населенного пункта
  «V248» #достигнутый уровень образования
)
data<- data[vars]
rm(vars)
#Зависимая переменная:
data$trustknown<- as.numeric(data$V104)*(-1)+5
data$trustunknown<- as.numeric(data$V105)*(-1)+5
data$index<- data$trustknown + data$trustunknown
#Независимые переменные:
data$income<- as.numeric(data$V239) #шкала доходов
data$gndr<- as.factor(data$V240) #пол
data$ages<- scale(as.numeric(data$V242)) #возраст
data$educ<- rep(NA, length(data$V248)) #уровень образования
```

¹ Код в электронной форме и используемая база данных по странам доступны по ссылке: <https://github.com/shirokaner/sociology4m>.

```
data$educ[data$V248 == "No formal education" |
data$V248 == "Incomplete primary school" |
data$V248 == "Complete primary school"] <- "primary"
data$educ[data$V248=="Incomplete secondary school: technical/
vocational type" |
data$V248 == "Complete secondary school: university-preparatory type" |
data$V248 == "Complete secondary school: technical/ vocational type" |
data$V248 == "Incomplete secondary school: university-preparatory type"
] <- "secondary"
data$educ[data$V248 == "Some university-level education, without
degree" |
data$V248 == "University - level education, with degree" ] <- "tertiary"
data$educ<- as.factor(data$educ)
data$V2 <- droplevels(data$V2)
names(data)[1] <- "country"
data<- merge(data, datac, all.x = TRUE)
data$logGDP<- scale(log(as.numeric(data$GDP)))
savevars<- c("country", "trustknown", "trustunknown", "index", "income",
"gn dr", "ages", "educf", "logGDP")
data1 <- na.omit(data[savevars] )
#Модели
m0 <- lmer(index ~ (1 | country), REML = FALSE, data = data1)
summary(m0) #80421 наблюдений, 56 стран
0.2039 / (0.2039 + 1.5685) # ICC = 11.5%
m1 <- lmer(index ~ ages + (1 | country), REML = FALSE, data = data1)
m2 <- lmer(index ~ gn dr + (1 | country), REML = FALSE, data = data1)
m3 <- lmer(index ~ income + (1 | country), REML = FALSE, data = data1)
m4 <- lmer(index ~ educf + (1 | country), REML = FALSE, data = data1)
anova(m0, m1)
anova(m0, m2)
anova(m0, m3)
anova(m0, m4)
m5 <- lmer(index ~ educf + gn dr + income + ages + (1 | country), REML =
FALSE, data = data1)
m6 <- lmer(index ~ educf + gn dr + income + ages + (1 + educf | country),
REML = FALSE, data = data1)
```

```
m7 <- lmer(index ~ gndr + ages + income + educf + logGDP + (1 + educf |
country), REML = FALSE, data = data1)
data1$educf5 <- relevel(data1$educf, ref = "secondary")
m7.1 <- lmer(index ~ gndr + ages + income + educf5 + logGDP + (1 +
educf5 | country), REML = F, data = data1)
m8 <- lmer(index ~ gndr + ages + income + educf * logGDP + (1 + educf |
country), REML = F, data = data1)
sjt.lmer(m1, m2, m3, m4, m5, p.kr = F, show.ci = FALSE, show.se =
TRUE, digits.est = 3)
sjt.lmer(m5, m6, m7, m8, p.kr = FALSE, show.ci = FALSE, show.se =
TRUE, digits.est = 3)
sjp.lmer(m0, sort.est = "(Intercept)", fade.ns = TRUE) #нулевая модель,
средние значения доверия по странам
sjp.lmer(m5, sort.est = «(Intercept)», fade.ns = TRUE, geom.colors =
c(«black», «black»)) #модель 5, means as outcomes
sjp.lmer(m8, sort.est = «(Intercept)», fade.ns = TRUE, geom.colors =
c(«black», «black»)) #модель 8, межуровневое взаимодействие, случай-
ный/рандомизированный эффект образования
sjp.int(m8, type = "cond", p.kr = FALSE) #маргинальный эффект обра-
зования на доверие
sjp.int(m8, type = "eff", p.kr = FALSE, show.ci = TRUE) #межуровневое
взаимодействие
#Диагностика модели
countrytrust<- tapply(data1$index, data1$country, mean) #визуальная
оценка
countrytrust<- as.data.frame(countrytrust)
countrytrust$country<- rownames(countrytrust)
datac1 <- merge(datac, countrytrust, all.x = TRUE)
datac1 <- na.omit(datac1)
ggplot(datac1) +
geom_point(aes(log(datac1$GDP), datac1$countrytrust), color = 'black') +
geom_text_repel(aes(log(datac1$GDP), datac1$countrytrust, label =
datac1$country.code3)) +
theme_classic(base_size = 16) + labs(x = «lnGDP PPP», y = «Trust
Index»)
m8i<- influence(m8, group = "country") #комплексная диагностика
```

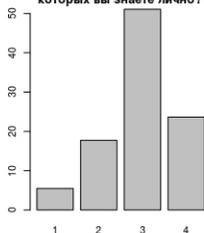
итоговой модели

```
cooktable<- cooks.distance(m8i) #расстояния Кука  
plot(m8i, which = «cook», cutoff = 4/56, sort = T, xlab = «Cook's Dis-  
tance», ylab = «Country»)  
dfetable<- dfbetas(m8i) #DFBETAs  
plot(m8i, which = «dfbetas», cutoff = 2/sqrt(56), xlab = «Cook's Distance»,  
ylab = «Country»)
```

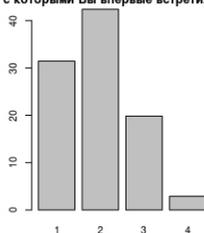
Приложение 2

КОМПОНЕНТЫ МОДЕЛИРУЕМОЙ ПЕРЕМЕННОЙ И ИХ РАСПРЕДЕЛЕНИЕ

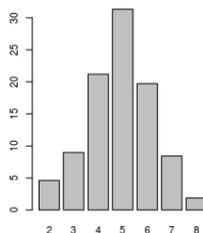
Распределение ответов на вопрос
"Насколько вы доверяете людям,
которых вы знаете лично?"



Распределение ответов на вопрос
"Насколько вы доверяете людям,
с которыми Вы впервые встретились?"

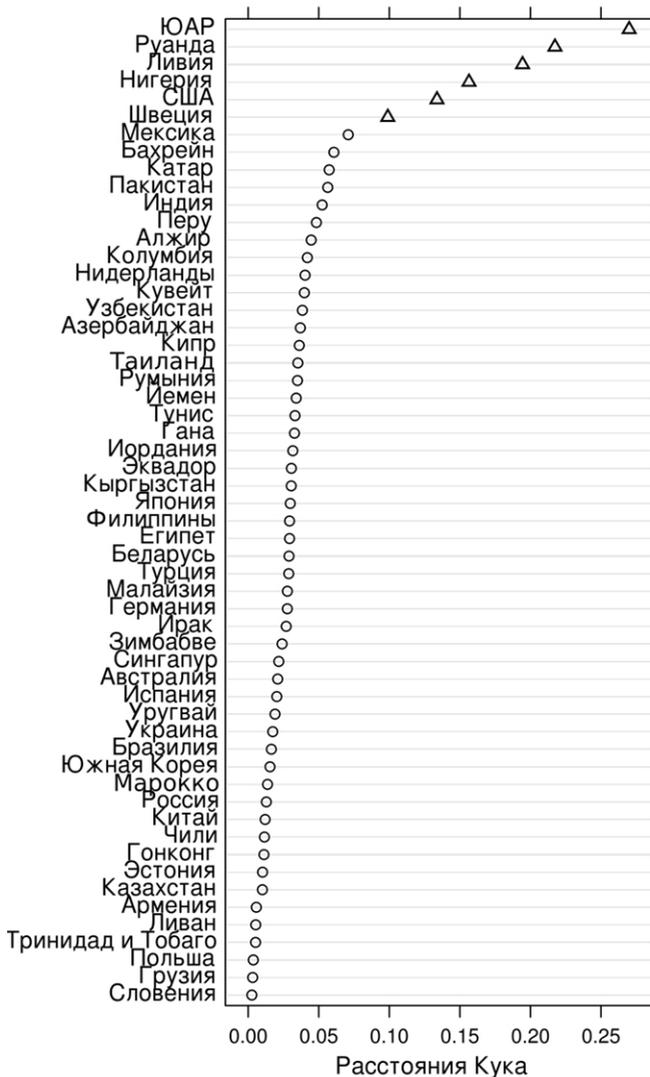


Распределение индекса доверия



Приложение 3

ДИАГНОСТИКА РЕГРЕССИИ: РАССТОЯНИЯ КУКА ПО СТРАНАМ



Volchenko Olesya

National Research University Higher School of Economics (NRU HSE),
Moscow, Saint Petersburg
ovolchenko@hse.ru

Shirokanova Anna

National Research University Higher School of Economics (NRU HSE),
Moscow, Saint Petersburg
ashirokanova@hse.ru

Applying multilevel regression modeling to cross-national data (on the example of generalized trust)

The paper deals with multilevel regression modeling (MLM) as a method preferred to the ordinary least-squares regression in the analysis of comparative data with hierarchical data structure. We present substantive reasons (contextual sources of heterogeneity, causal heterogeneity, and generalizability of results) and statistical reasons (obtaining more precise and reliable estimates) for multilevel modeling. We also provide an overview of MLM implementation in several statistical packages. Using the cross-national World Values Survey (WVS) data, we outline a step-by-step procedure for building and fitting a two-level linear regression model of generalized trust on educational attainment levels (the “null” model, the fixed-intercept model, the random-intercept model, the random-intercept random-slope model, the model with a country-level predictor, and the cross-level interaction model). Then we describe and compare existing goodness-of-fit measures for MLM (AIC, BIC, maximum likelihood functions, and pseudo- R^2). We also demonstrate robustness check techniques for multilevel models (visualization, Cook’s distance, and DFBETAs). In the final section, we overview alternative approaches to multilevel modeling when dealing with hierarchical data (cluster robust standard errors, generalized estimating equations, country fixed effects, country means, and aggregation) as currently practiced in comparative cross-national social science research. The replicable R code is attached.

Key words: comparative research, hierarchical models, multilevel modeling, WVS, World Values Survey, trust, R statistical environment, regression analysis, cross-national data

References

1. Raudenbush S.W., Bryk A.S. *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks: Sage, 2002.
2. Gelman A., Hill J. *Data analysis using regression and multilevel/hierarchical models*. New York: Cambridge University Press, 2006.
3. Mills M.C., Präg P. “Methodological advances in cross-national research: Multilevel challenges and solutions”, *European sociological review*, 2016, 1, 1–2.
4. McNeish D., Stapleton L.M., Silverman R.D. “On the Unnecessary Ubiquity of Hierarchical Linear Modeling”, *Psychological Methods*, 2016. URL: <http://dx.doi.org/10.1037/met0000078> (date of access: 30.11.2016)
5. Rudnev M.G. “The influence of the country of origin and the country of residence on basic values (results of a study of European migrants)” (in Russian), in: *XIII April International Academic Conference on Economic and Social Development*, 2012, vol. 2, p. 591-604. URL: <https://publications.hse.ru/chapters/67471834> (date of access: 30.04.2017)
6. Salnikova D.V. “Modelling the relationship between subjective economic well-being of citizens and welfare state institutions in the EU countries” (in Russian), *Applied Econometrics (Prikladnaya ekonometrika)*, 2014, 1, 71-89.
7. Rusinova N.L., Safronov V.V. “Health in the European States: Status Inequalities and Social Capital” (in Russian), *Sociological Journal (Sotsiologicheskij Zhurnal)*, 2015, 21 (4), 34-60.
8. Kostenko V.V. “Gender attitudes of Muslim migrants in Western Europe: multilevel non-hierarchical analysis” (in Russian), *Monitoring of Public Opinion: Economic and Social Changes*, 2017, 1, 97-120.
9. Reisinger W.M., Pacheco J. “Mass Democratic Values and Democracy in Russia’s Regions” (in Russian), *Monitoring of Public Opinion: Economic and Social Changes*, 2014, 5, 12-30.
10. Sedova E.N., Stebunova O.I., Ushatova S.T. “Econometric modelling of residential property prices at regional level: hierarchical approach” (in Russian), *Intelligence. Innovations. Investments (Intellekt. Innovacii. Investicii)*, 2016, 4, 51-56.
11. Titkova V.V., Ivanushina V.A., Alexandrov D.A. “Pupils’ popularity and an educational setting at school” (in Russian), *Educational Studies (Voprosy Obrazovaniya)*, 2013, 4, 145-167.
12. Uvarov A.G., Yastrebov G.A. “Schools and socioeconomic standing of Russian families as competing factors in promoting social inequality in Russia” (in Russian), *Universe of Russia. Sociology. Ethnology (Mir Rossii. Sociologiya. Etnologiya)*, 2014, 23 (2), 103-132.
13. Kuzmina Y.V. “Big-frog-in-a-small-pond effect: Is it always good for a child to study in a strong class?” (in Russian), *Psychology. Journal of the Higher School of Economics (Psichologiya. Zhurnal vysshej shkoly ekonomiki)*, 2016, 13(4), 712-740.

14. Chetverikov A.A. “Linear Mixed Effects Regression in Cognitive Studies” (in Russian), *Russian Journal of Cognitive Science (Rossijskij zhurnal kognitivnoj nauki)*, 2015, 2(1), 41-51.
15. Alexeeva S.V., Slioussar N.A. “Parafoveal processing in reading: the role of word length” (in Russian), *Tomsk State University Journal of Philology (Vestnik Tomskogo gosudarstvennogo universiteta Filologiya)*, 2017, 45, 5-29.
16. Luke D.A. *Multilevel modelling*. Thousand Oaks: Sage, 2004.
17. Snijders T.A.B., Bosker R.J. *Multilevel analysis: An introduction to basic and advanced multilevel*. Thousand Oaks: Sage, 2012.
18. Kreft I.G.G., Leeuw J. de. *Introducing multilevel modelling*. London: Sage, 1998.
19. Spencer N.H. “Models for value-added investigations of teaching styles data”, *Journal of data science*, 2008, 1, 33–51.
20. Geiser C. *Data analysis with Mplus*. New York: Guilford Press, 2013.
21. Bryan M.L., Jenkins S.P. “Multilevel modeling of country effects: A cautionary tale”, *European sociological review*, 2016, 1, 3–22.
22. Subramanian S.V. et al. “Revisiting Robinson: The perils of individualistic and ecologic fallacy”, *International journal of epidemiology*, 2009, 2, 342–360.
23. Bartels B. “Beyond ‘fixed versus random effects’: A framework for improving substantive and statistical analysis of panel, time-series cross-sectional, and multilevel data”, in: *Quantitative research in political science*, Thousand Oaks: SAGE Publications, 2008.
24. World Values Survey Association. *World values survey wave 6 2010-2014 official aggregate.v.20150418.Asep/JDS [producer]*, 2016. URL: <http://www.worldvaluessurvey.org/WVSDocumentationWVL.jsp> (date of access: 30.09.2016)
25. Delhey J., Newton K. “Who trusts?: The origins of social trust in seven societies”, *European societies*, 2003, 2, 93–137.
26. Hamamura T. “Social class predicts generalized trust but only in wealthy societies”, *Journal of cross-cultural psychology*, 2012, 3, 498-509.
27. Charron N., Rothstein B. “Does education lead to higher generalized trust? The importance of quality of government”, *International journal of educational development*, 2016, 50, 59–73.
28. Frederiksen M., Larsen C.A., Lolle H.L. “Education and trust Exploring the association across social relationships and nations”, *Acta Sociologica*, 2016, 59 (4), 294-308.
29. Diez Roux A.V. “A glossary for multilevel analysis”, *Journal of epidemiology and community health*, 2002, 56 (8), 588-594.
30. Fairbrother M. “Two Multilevel Modeling Techniques for Analyzing Comparative Longitudinal Survey Datasets”, *Political Science Research and Methods*, 2014, 2 (1), 119–140.

31. Bell A., Jones K. “Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data”, *Political Science Research and Methods*, 2015, 3 (1), 133-153.
32. Mewes J. “Gen(d)eralized trust: Women, work, and trust in strangers”, *European sociological review*, 2014, 3, 373–386.
33. Steenbergen M.R., Jones B.S. “Modeling multilevel data structures”, *American journal of political science*, 2002, 1, 218–237.
34. Gill J., Womack A.J. “The multilevel model framework”, in: *The SAGE handbook of multilevel modeling*. Thousand Oaks: SAGE Publications, 2013. P. 3–20.
35. Tate R.L., Wongbunhit Y. “Random versus nonrandom coefficient models for multilevel analysis”, *Journal of educational and behavioral statistics*, 1983, 2, 103–120.
36. Hox J.J. “Multilevel modeling: When and why”, in: *Classification, data analysis, and data highways*. Berlin, Heidelberg: Springer, 1998. P. 147–154.
37. Maas C.J., Hox J.J. “Sufficient sample sizes for multilevel modelling”, *Methodology*, 2005, 3, 86–92.
38. Preacher K.J., Zhang Z., Zyphur M.J. “Alternative methods for assessing mediation in multilevel data: The advantages of multilevel SEM”, *Structural equation modeling*, 2011, 2, 161–182.
39. Muthén B. “Latent variable modeling of longitudinal and multilevel data”, *Sociological methodology*, 1997, 1, 453–480.
40. Erceg V. et al. “An empirically based path loss model for wireless channels in suburban environments”, *IEEE Journal on selected areas in communications*, 1999, 7, 1205–1211.
41. Austin P.C., Goel V., Walraven C. van “An introduction to multilevel regression models”, *Canadian journal of public health*, 2001, 2, 150-154.
42. Hofmann D.A. “An overview of the logic and rationale of hierarchical linear models”, *Journal of management*, 1997, 23 (6), 723–744.
43. Barr D.J. et al. “Random effects structure for confirmatory hypothesis testing: Keep it maximal”, *Journal of memory and language*, 2013, 3, 255–278.
44. Townsend Z. et al. “The choice between fixed and random effects”, in: *The SAGE handbook of multilevel modeling*. Thousand Oaks: SAGE Publications, 2013. P. 73–88.
45. Schmidt-Catran A.W., Fairbrother M. “The random effects in multilevel models: getting them wrong and getting them right”, *European sociological review*, 2015, 1, 23–38.
46. Gelman A., Pardoe I. “Bayesian measures of explained variance and pooling in multilevel (hierarchical) models”, *Technometrics*, 2006, 2, 241–251.
47. Xu R. “Measuring explained variation in linear mixed effects models”, *Statistics in medicine*, 2003, 22, 3527–3541.

48. Orelien J.G., Edwards L.J. “Fixed-effect variable selection in linear mixed models using R2 statistics”, *Computational Statistics & Data Analysis*, 2008, 52, 1896–1907.
49. Hox J.J. *Multilevel analysis: Techniques and applications*. Hove: Routledge, 2010.
50. Meer T. van der, Grotenhuis M.T., Pelzer B. “Influential cases in multilevel modeling: A methodological comment”, *American sociological review*, 2010, 1, 173–178.
51. Aguinis H., Gottfredson R.K., Joo H. “Best-Practice Recommendations for Defining, Identifying, and Handling Outliers”, *Organizational Research Management*, 2013, 16 (2), 270–301.
52. Nieuwenhuis R., Grotenhuis H.F. te, Pelzer B.J. “influence.ME: Tools for detecting influential data in mixed effects models”, *The R-Journal*, 2012, 2, 38–47.
53. Hayes A.F. “A primer on multilevel modeling”, *Human communication research*, 2006, 4, 385–410.
54. Delhey J., Welzel C. “Generalizing trust: What extends trust from ingroups to outgroups?”, *World values research*, 2012, 3, 45–69.
55. Gheorghiu M.A., Vignoles V.L., Smith P.B. “Beyond the United States and Japan: Testing Yamagishi’s emancipation theory of trust across 31 nations”, *Social psychology quarterly*, 2009, 4, 365–383.
56. Robbins B.G. “Neither government nor community alone: A test of state-centered models of generalized trust”, *Rationality and society*, 2011, 3, 304–346.
57. Almakaeva A. M. “Measurement of generalized trust in cross-cultural studies” (in Russian), *Sotsiologicheskie issledovaniya (Sociological studies)*, 2014, 11, 32–43.
58. Alesina A., La Ferrara E. “Who trusts others?”, *Journal of public economics*, 2002, 2, 207–234.
59. Yamagishi T. “Trust as a form of social intelligence”, in: *Trust in society*. New York: Russell Sage Foundation, 2001. P. 121–147.
60. Keefer P., Knack S. “Social capital, social norms and the new institutional economics”, in: *Handbook of new institutional economics*. Berlin, Heidelberg: Springer, 2008. P. 701–725.
61. Helliwell J.F., Putnam R.D. “Education and social capital”, *Eastern economic journal*, 2007, 1, 1–19.
62. Fabrikant M. S. “A model-oriented approach to missing values: multiple imputation in multilevel regression using R (poll data case)” (in Russian), *Sotsiologiya 4M (Sociology: methodology, methods, mathematical modeling)*, 2015, 41, 7–29.
63. Petersen M.A. “Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches”, *Review of Financial Studies*, 2009, 22, 435–480.

64. Primo D.M., Jacobsmeier M.L., Milyo J. “Estimating the impact of state policies and institutions with mixed-level data”, *State Politics & Policy Quarterly*, 2007, 4, 446–459.
65. Gelman A. “Multilevel (hierarchical) modeling: what it can and cannot do”, *Technometrics*, 2006, 48 (3), 432-435.
66. Rue H., Martino S., Chopin N. “Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations”, *Journal of the royal statistical society: Series b (statistical methodology)*, 2009, 71 (2), 319–392.