

5. МЕТОДИЧЕСКИЕ ЭКСПЕРИМЕНТЫ И НОВАЦИИ

Эмпирическое сравнение статистических алгоритмов обнаружения фальсификаций в опросах потребителей

Бабич Николай Сергеевич,
ООО «Эксперткон»

Немец Хайнер Доррох в течение 19 лет трудился интервьюером в сетях сразу нескольких крупнейших опросных центров ФРГ. Человек аккуратный, неглупый и даже творческий, он зарабатывал на жизнь себе и своей семье, выполняя солидный объем работ для Ifak, Emnid (сейчас TNS-Emnid), GfM-Getas (сейчас Ipsos Germany), GfK и других компаний. Все это время Доррох фальсифицировал анкеты, обманывая своих нанимателей не только под собственным именем, но и под несколькими псевдонимами. О масштабах деятельности предприимчивого отца семейства стало известно только после того, как он написал книгу о своем большом опыте¹.

Приведенный пример, конечно, нетипичен, но возникает естественный вопрос о том, насколько он исключителен, и какие меры могут служить адекватным ответом на эту угрозу. Немецкие социологи, реагируя на книгу Дорроха, приводили оценки доли фальсификаций в диапазоне от менее трех процентов до трети всех анкет². Представляется, что называвшаяся верхняя планка слишком пессимистична, но даже 3 % фальсификаций являются серьезным ударом по качеству данных. Интуитивно эта доля кажется небольшой, но необходимо помнить, что для снижения предельной погрешности на 3 %, простую случайную выборку увеличивают примерно с 400 человек до 2000, то есть в пять раз (предполагая бесконечную генеральную совокупность и доверительную вероятность 0,95). Кроме того, поддельные анкеты могут оказывать серьезное влияние на результаты многомерных методов анализа, так как из-за неспособности фальсификаторов воспроизвести существующие в данных связи существенно смещаются коэффициенты корреляции³. Все это свидетельствует о высокой актуальности развития системы контроля достоверности в массовых опросах.

Контроль достоверности, как органичная часть процесса контроля качества, сам является комплексной задачей, решение которой предполагает разработку системы адекватной мотивации интервьюеров, их обучение и профессиональную социализацию, тщательную апробацию анкеты, профилактическое включение в выборку ложных адресов и т.п. Но наиболее затратной частью контроля достоверности выступает повторный опрос, то есть, эмпирический контроль работы интервьюеров. Повторный опрос, в конечном итоге, — одно из немногих твердых оснований для принятия кадровых решений и финансовых санкций в полевых работах. Поэтому большой интерес представляют методы оптимизации контрольной выборки, позволяющие в первую очередь проверять именно те анкеты, которые имеют наибольшие шансы оказаться сфабрированными. Статистические алгоритмы обнаружения фальсификаций работают именно по этому принципу, они не могут сами по себе служить доказательствами недобросовестной работы интервьюеров, но должны играть роль в выборе объектов контроля. Такая постановка задачи выводит нас на очевидный критерий сравнения эффективности указанных алгоритмов, а именно: долю сфабрированных анкет в формируемой с их помощью контрольной выборке.

¹ Dorroch H. Meinungsmacher-Report. Göttingen, 1994.

² Ноэль-Нойман Э., Петерсен Т. Все, но не каждый. М., 2007. С. 360; Ohrfeige an der Haustür // Der Spiegel. 1994. №26. S. 41.

³ Schnell R. Der Einfluss gefälschter Interviews auf Survey-Ergebnisse // Zeitschrift für Soziologie. 1991. №1. S. 25–35; Schräpler J. P., Wagner G. Characteristics and impact of faked interviews in surveys: an analysis of genuine fakes in the raw data of SOEP // Allgemeines Statistisches Archiv. 2005. №1. P. 7–20.

Точкой отсчета, с которой можно сравнивать эту эффективность, является та же доля в выборке случайной.

Алгоритмы обнаружения фальсификаций

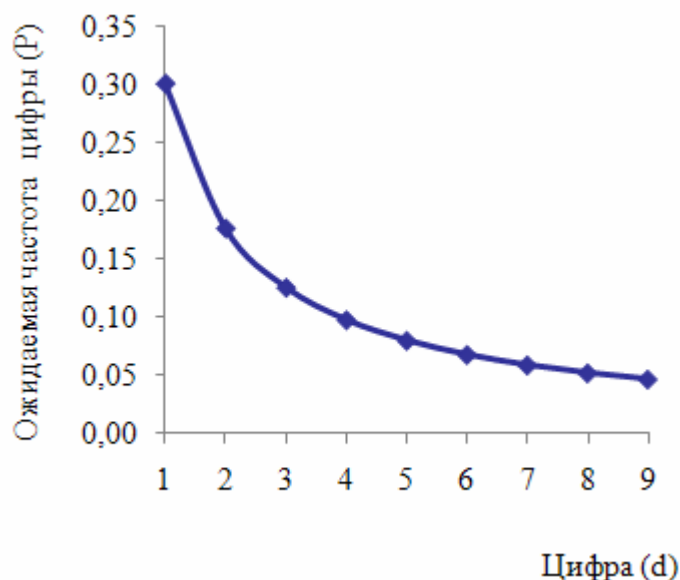
Обобщение литературы и собственного практического опыта позволяет нам говорить о существовании как минимум четырех статистических алгоритмов, которые позволяют выявить подозрительные на предмет фальсификации анкеты в массиве данных опроса.

Первый из алгоритмов, вероятно, наиболее распространенный, основан на так называемом «законе Бенфорда»¹. Суть его состоит в том, что первые цифры в значениях некоторых социальных переменных встречаются с определенной вероятностью, которая тем меньше, чем больше цифра. Эта закономерность может проявляться при соблюдении равных условных «затрат» на получение каждой единицы измеряемого показателя, в процессах, характеризующихся экспоненциальным ростом, и при ряде некоторых других специальных условий. Очень трудно априорно установить соответствие социологической переменной закону Бенфорда, однако во многих случаях можно обоснованно его предполагать.

Рассмотрим упрощенный пример с доходом респондента. Чтобы заработать тысячу рублей, человек прикладывает определенные усилия. Для получения двух тысяч ему требуется уже в два раза больше усилий. Но чтобы заработать три тысячи, он должен потрудиться только в полтора раза больше, чем ради двух тысяч. Если предположение о равных удельных затратах труда работает, это означает, что вероятность встретить в качестве первого знака в оценке дохода единицу наиболее высока. Двойка встретится в качестве первого знака в этой оценке значительно реже, а тройка — еще реже, но не намного реже, чем двойка. Таким образом, частота цифр, встречающаяся в вопросе о доходах, складывается в математически описываемую последовательность, которая и известна как «закон Бенфорда»:

$$P_d = \lg \left(1 + \frac{1}{d} \right)$$

В приведенной формуле d — это первая цифра в значениях некоторой переменной, P_d — вероятность появления этой цифры. Закон Бенфорда имеет и достаточно простое графическое выражение (рис. 1).



Отклонение от приведенного распределения может сигнализировать о том, что данные были сфальсифицированы. Не исключено, что некоторые интервьюеры интуитивно чувствуют, какова приблизительно должна быть частота тех или иных цифр. Но ее воспроизведение в соответ-

¹ См., например: Swanson D., Cho M. J., Eltinge J. Detecting Possibly Fraudulent or Error-Prone Survey Data Using Benford's Law // American Statistical Association: Proceedings of The Section on Survey Research Methods. 2003. P. 4172–4177; Wang Y., Pedlow S. Detecting Falsified Cases in SCF 2004 Using Benford's Law // Proceedings of the Survey Research Methods Section of the American Statistical Association. Alexandria VA, 2005. P. 3652–3657.

ствии с законом Бенфорда без специальной математической подготовки маловероятно. Естественной реализацией данного принципа проверки является последовательное исключение (с возвращением) всех анкет массива, сопровождающееся вычислением соответствия массива закону Бенфорда для каждого такого исключения. Те анкеты, исключение которых *повышает* соответствие эмпирического распределения теоретическому, должны быть признаны подозрительными. Если анализ проводится по нескольким признакам, результат по каждому из них можно считать вкладом в общую сумму «штрафных очков» анкеты.

Второй алгоритм, который мы условно обозначим как «дисперсионный», основан на следующих простых соображениях относительно логики работы недобросовестного интервьюера¹. Всякому здравомыслящему фальсификатору ясно, что при фабрикации ответов по оценочным шкалам (например, по психографическим инструментам, шкалам оценки продуктов и т.п.) он должен стараться избегать крайних вариантов. Крайние варианты приводят к повышению риска обнаружения фабрикации, так как при проверке они легче опровергаются. Но, стремясь избежать крайних вариантов, фальсификатор попадает в статистическую ловушку, так как он снижает вариативность своих данных по сравнению с выборкой в целом. Следовательно, если у какого-то интервьюера дисперсия меньше, чем во всей выборке, это значит, его анкетам надо уделить более пристальное внимание.

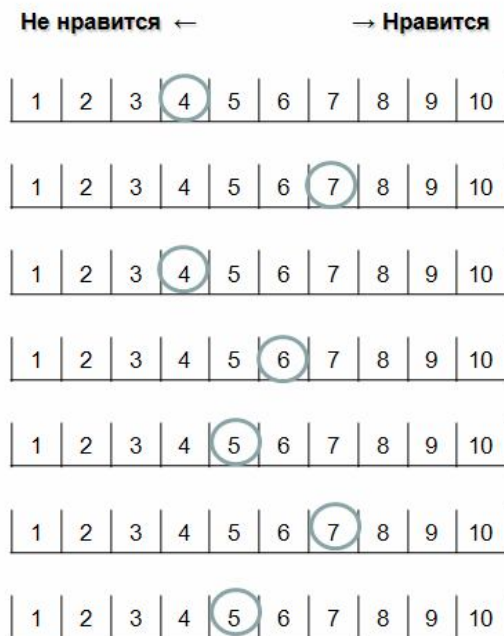


Рис. 2. Паттерн ответов недобросовестного интервьюера.

Алгоритм, работающий для каждой отдельной анкеты, аналогично предыдущему, вычисляет дисперсию при последовательном исключении данных каждого респондента. Результаты по нескольким переменным могут быть просуммированы, и общее число «штрафных очков», теоретически, должно быть более надежным показателем подозрительности анкеты с точки зрения фальсификации.

Третий алгоритм, «кластерный», основан на выделении типологической группировки, которая, в случае фальсификации, должна значительно совпадать с группировкой анкет по интервьюерам. Интервьюер, как и любой человек, обладает уникальным и в то же время ограниченным социальным опытом. Поэтому его ожидания часто образуют конфигурацию, сильно отличающуюся от реальности. Заполняя анкету за респондента, интервьюер представляет себе, что тот мог бы ответить. И этот процесс неизбежно основывается на стереотипах. Если мы определим какую-нибудь меру стереотипности базы данных опроса, то получим инструмент оценки ее

¹Schafer C., Schrapler J.-P., Muller K. R., Wagner G. G. Automatic Identification of Faked and Fraudulent Interviews in Surveys by Two Different Methods // American Statistical Association: Proceedings of the Section on Survey Research Methods. 2005. P. 4318–4325.

достоверности. В этой роли может выступить сила связи между интервьюером и типологической группировкой ответов.

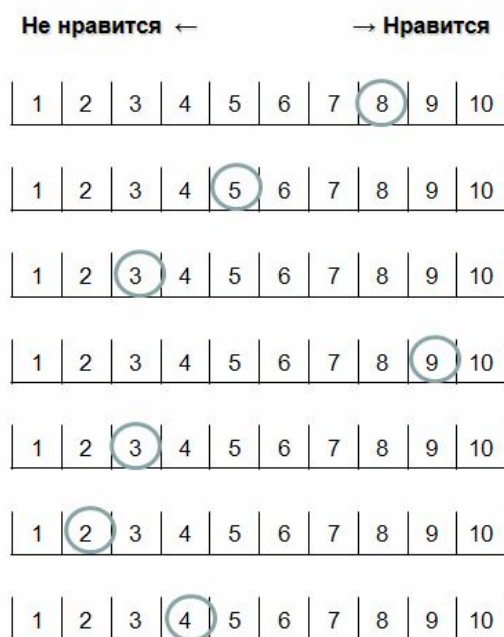


Рис. 3. Паттерн ответов добросовестного интервьюера.

Кластерный анализ позволяет типологизировать респондентов. Это первый этап проверки. На втором этапе следует построить таблицу сопряженности между переменными, нумерующими интервьюеров и типы ответов. Сильная связь между ними может означать высокую степень влияния интервьюерских стереотипов на данные. Оценка достоверности каждой анкеты происходит по той же схеме, только рассчитывается для каждого исключения сила связи между нумерацией интервьюеров и типологической группировкой ответов. Если мы наблюдаем ослабленные связи после исключения анкеты, ей должны быть присуждены «штрафные очки». Поскольку кластерный анализ — неустойчивый метод, типологизацию респондентов целесообразно проводить несколько раз, применяя определенную комбинацию алгоритмов объединения и мер расстояния. Начисленная по итогам сумма «штрафных очков» может трактоваться как мера подозрительности анкеты с точки зрения фальсификации.

Четвертый алгоритм базируется на статистических тестах для проверки гипотез о различиях между группами. Для простоты, мы будем рассматривать только случай проверки гипотезы о различиях между несколькими группами по метрическим переменным критерием Краскала–Уоллиса. Ясно, что существенные отличия интервьюера от всей выборки свидетельствуют о возможных фальсификациях. Каждая отдельная анкета в этом аспекте может быть проанализирована точно так же, как и в предыдущих случаях, но рассчитываемым при каждом исключении показателем будет уровень значимости различий между интервьюерами по критерию Краскала–Уоллиса. Снижение значимости различий будет маркировать анкету как подозрительную.

Дизайн эксперимента

Для эмпирической проверки эффективности четырех описанных алгоритмов был проведен методический эксперимент. Материалом для него послужили результаты опроса потребителей компьютерных игр (в возрасте от 15 до 45 лет), проведенного ООО «Эксперткон» в ноябре 2010 года методом личного формализованного интервью на выборке в 400 человек. Данные опроса были в высокой степени надежными, так как основой выборки послужила база данных программы лояльности заказчика, что позволило при проверке задать необходимые вопросы 84% респондентов (остальные по разным причинам оказались недоступны для повторного опроса). Анкеты, полностью аналогичные использованным в ноябре, месяцем позже были предложены для «фабрикации» четверем исследователям, имеющим опыт интервьюирования. Им не были сообщены результаты реального опроса, но было дано задание изобретать ответы таким образом, чтобы их как можно труднее было поймать.

Каждый исследователь сфабриковал 25 анкет, после чего общий массив достиг 500 наблюдений. К полученной базе данных были последовательно применены все четыре описанных выше алгоритма. Алгоритм, основанный на законе Бенфорда, применялся к переменным, описывающим доход респондента, срок знакомства с играми заказчика, число раз, которое респондент сталкивался с рекламой заказчика за три месяца, число покупок компьютерных игр за три месяца. Все остальные алгоритмы применялись к двум батареям шкал: психографической батарее, описывающей стремление к поиску новых ощущений (10 шкал) и батарее оценки товара заказчика по потребительским свойствам (8 шкал). Кластерный анализ выполнялся с использованием разных алгоритмов и разных мер расстояния 18 раз, после чего для каждой из 18 кластеризаций были посчитаны «штрафные очки». Общая сумма «штрафных очков» для алгоритма, основанного на законе Бенфорда, колебалась от 0 до 4, для всех остальных алгоритмов — от 0 до 18.

После проведения анализа массив был последовательно упорядочен по сумме штрафных очков каждого алгоритма, начиная от максимальной. Нас интересовал следующий вопрос. Если мы используем статистический анализ для формирования контрольной выборки, скажем, 20 % от всего массива, то насколько алгоритмы обнаружения фальсификаций улучшат эту выборку по сравнению со случайным распределением контрольных интервью? Иными словами, сколько будет сфабрикованных анкет в первых 100, упорядоченных по штрафным очкам? Если в качестве эталона рассматривать случайную выборку, то доля фальсификата, естественно, должна стремиться к 20 %, с доверительным интервалом 7 % для доверительной вероятности 0,95 и 9 % для доверительной вероятности 0,99. Таким образом, если описанные алгоритмы дают от 30 % и более в 100 наиболее оштрафованных анкетах, они работают лучше, чем случайный отбор.

Обсуждение результатов

Основные итоги эксперимента представлены на рис. 4. Как следует из диаграммы, наименее благоприятными они оказались для закона Бенфорда. Хотя результаты его использования остаются лучше случайного отбора, они существенно уступают остальным трем алгоритмам. Из 100 наиболее подозрительных анкет, только 34 оказались сфабрикованными. Иными словами, критерий Бенфорда позволил лишь немного улучшить результат случайного отбора. Это может быть объяснено тем, что данный алгоритм обнаружения фальсификаций применялся только на четырех переменных, а не на восемнадцати, как его конкуренты. Тем самым был уменьшен объем информации, задействованный в анализе, и алгоритм оказался в заведомо проигрышной ситуации. Но такое «неравноправие» отражает объективную ситуацию, так как только эти четыре переменные и подходили для его применения во всем массиве.

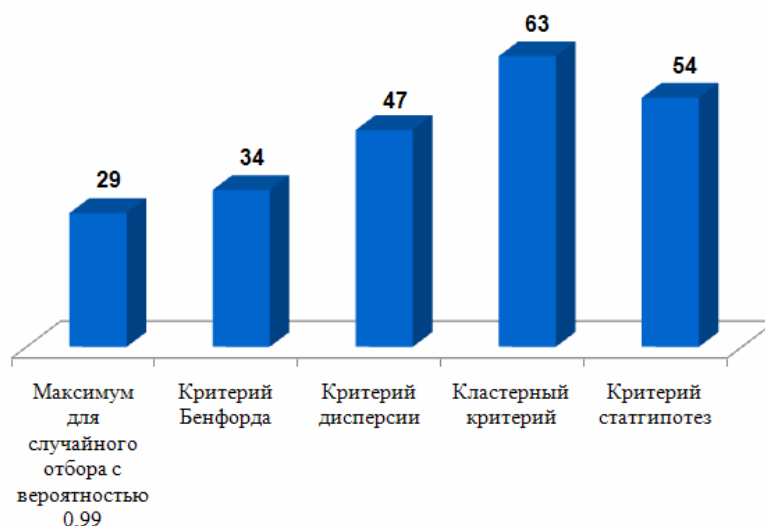


Рис. 4. Число подделок в первой сотне анкет.

Чтобы закон Бенфорда выполнялся, требуется соответствие действительности достаточно сильных предположений, лежащих в его основе. Априорно допустить правомерность этих предположений можно лишь для очень небольшого числа шкал, встречающихся в социологиче-

ских и маркетинговых исследованиях. Поэтому на практике применение критерия Бенфорда при поиске фальсификаций всегда будет сталкиваться с недостатком данных.

«Дисперсионный» алгоритм и алгоритм, основанный на проверке статистических гипотез, дали существенно лучший результат — 47 и 54 сфабрикованных анкеты, соответственно. Абсолютным лидером по эффективности оказался кластерный алгоритм, который позволил сосредоточить в контрольной выборке 63 сфабрикованные анкеты из 100. Большой успех кластерного критерия, возможно, объясняется более интенсивным использованием имеющейся в наличии информации, так как в нем 18 раз учитывались все 18 шкал, включенных в анализ.

Представляет интерес распределение результатов критериев по интервьюерам (таблица). Первое, на что необходимо обратить внимание — попадание в контрольную выборку всех «недобросовестных интервьюеров» при использовании каждого из алгоритмов. В этом смысле использованные критерии достаточно качественно решают свою задачу. Второй важный момент — неравномерная успешность разных алгоритмов применительно к разным интервьюерам. За исключением, быть может, кластерного критерия, во всех остальных случаях наблюдается более и менее удачная для фальсификатора стратегия обмана алгоритма. Вероятно, что критерий дисперсии хуже работает там, где лучше работает критерий статистических гипотез, и наоборот. Это вполне правдоподобно, так как первый основан на разоблачении попыток имитировать «усредненного» респондента, а второй — на различиях средних (или, как в данном случае, медиан).

Таблица

Распределение результативности критериев по интервьюерам

Интервьюер	Критерий Бенфорда	Критерий дисперсии	Кластерный критерий	Критерий статгипотез
1	5	15	14	8
2	6	19	18	11
3	13	6	16	16
4	10	7	15	19
Итого	34	47	63	54

Нельзя не отметить, что полученные результаты имеют существенные ограничения. Во-первых, они получены в искусственной ситуации, которая не может претендовать на полную имитацию реальных процессов, происходящих при опросе. Во-вторых, в эксперимент было вовлечено сравнительно небольшое число участников, что делает всякие сравнения между стратегиями фальсификации весьма предположительными. В-третьих, рассмотренные и оцененные алгоритмы представляют собой во многих аспектах достаточно сложные процессы, течение которых однозначно предсказать невозможно. Особенно это касается кластерного анализа. Тем не менее, проведенное исследование во многих отношениях показательное и позволяет выдвинуть ряд практических рекомендаций, а также предположить, какие направления будущей работы могут оказаться наиболее плодотворными.

Заключение

Из четырех алгоритмов, которые сравнивались в проведенном исследовании, наименее предпочтительным выглядит основанный на законе Бенфорда. Он неудобен для использования в массовых опросах в силу своих ограничений, а для тех переменных, к которым его, теоретически, можно применить, этот алгоритм дает неоднозначные результаты. По-видимому, должны объясняться скорее успешные случаи его использования, чем неудачи. Тем не менее, и этот алгоритм показывает результаты лучше, чем можно ожидать при случайном отборе.

Наиболее многообещающим представляется метод выявления фальсификаций, основанный на кластерном анализе. Он продемонстрировал свою высокую эффективность в проведенном испытании, хотя, из-за естественных ограничений генерализации любого методического эксперимента, нельзя утверждать, что именно этот алгоритм будет наилучшим во всех случаях. Но именно кластерный критерий содержит в себе наибольший потенциал для усовершенствования, так как использованные в данный момент в комбинации методы кластерного анализа могут

быть разделены, с последующим отбором и теоретическим обоснованием оптимальных процедур. В связи с различиями эффективности рассмотренных критериев, применяемых к разным интервьюерам, также возникает вопрос о создании их наилучшей комбинации.

Критерий дисперсии также продемонстрировал положительные результаты, но его дальнейшее широкое использование мы не можем рекомендовать. Этот критерий — единственный, который основан на легко воспроизводимых в процессе фальсификации моделях данных. При проникновении информации о нем в среду интервьюеров, недобросовестным сотрудникам не составит никакого труда обойти соответствующий фильтр.

Алгоритм, основанный на проверке статистических гипотез, представляет большой интерес для дальнейших исследований по двум причинам. Во-первых, он демонстрирует сравнительно высокую эффективность. Во-вторых, математическая дисциплина проверки статистических гипотез достаточно развита, чтобы здесь представлялось более вероятным, чем в кластерном анализе, детальное теоретическое обоснование совокупности возможных алгоритмов и выбор их оптимальной конфигурации.

Все четыре рассмотренных алгоритма для своего применения требуют программной реализации. Она не слишком сложна, и поэтому формирование эффективной выборки для контроля, в принципе, возможно с использованием обычных инструментов синтаксиса SPSS, макросов Excel и Statistica. Но логика развития метода все же требует, чтобы итогом изысканий во всех отмеченных направлениях стало программное обеспечение, которое позволит полевым отделам исследовательских агентств применять полученные результаты в потоковом режиме. Разработка такого программного продукта в настоящее время является одной из наших приоритетных задач.

Влияние метода сбора данных на искренность ответов на чувствительные вопросы¹

Григорьева Марина Валерьевна,
*Ивановский государственный энергетический
университет им. В.И. Ленина*

В 2009 году на выборке студентов ивановских вузов (N=300) изучалось, как влияет вид анкетирования на ответы респондентов о рискованном сексуальном поведении. Сравнивались персональное раздаточное анкетирование, групповое раздаточное анкетирование и компьютерное персональное анкетирование в группе. Мы ожидали (гипотеза), что компьютерный опрос на чувствительные темы повысит искренность ответов: устранил влияние присутствия другого, снизит эффект интервьюера, снизит пропуск ответов и др. В данном случае под искренностью мы понимаем «соответствие между тем, что человек хотел сказать, и тем, что он думает об этом на самом деле»²; неискренность возникает в результате редактирования респондентами своих «истинных» ответов с учетом идеального образа «Я», жизненных целей, мотивов и интересов³. Таким образом, повышение доли социально нежелательных ответов при изменении метода опроса мы трактуем как положительное влияние данного метода на искренность опрашиваемых. В целом, анализ показал, что влияние метода опроса на искренность ответов респондентов о сексуальном поведении опосредуется не/одобрительным отношением респондентов к соответствующему виду рискованного поведения.

Для измерения отношения респондентов к отдельным видам сексуального поведения использовалась разработанная автором методика, позволяющая получить сводный индекс отношения к тем или иным видам поведения. Для расчета данного индекса по каждому виду сексуального поведения задавался вопрос типа: «На Ваш взгляд, иметь сексуальные контакты со случайными партнерами... допустимо; допустимо, но не желательно; не допустимо?».

¹ Автор выражает благодарность своей бывшей студентке Е.А. Дождиковой за сотрудничество при проведении данного исследования.

² Кожевникова О.В., Кулагина Е.А. Искренность в ответах респондентов о политических предпочтениях // Мониторинг общественного мнения. 2008. № 2. С. 38.

³ Мягков А.Ю. Всегда ли респонденты говорят правду? Мета-анализ зарубежных источников // Социологические исследования. 2008. № 9. С. 24.

Подсчет средних проводился по 10-балльной шкале по трем группам респондентов: (1) считающих, что предложенные виды сексуальных практик допустимы, (2) допустимы, но не желательны, (3) не допустимы.

На основе средних мы сконструировали пять шкал-континуумов, которые отражали показатели отношения к (А) сексуальным контактам со случайными партнерами, (Б) сексуальным контактам в нетрезвом состоянии, (В) большому числу сексуальных партнеров, (Г) сексуальным контактам со случайными партнерами при наличии постоянного партнера, (Д) сексуальным контактам с партнером своего пола. Для корректности сравнения показателей отношения к исследуемым видам сексуальных практик за точку отсчета на этих шкалах мы приняли средние оценки для категории респондентов, считающих, что указанные виды сексуальных отношений допустимы, но не желательны, то есть тех, кто занимает в целом нейтральные позиции. Все значения средних, попавшие в область ниже заданного критерия (левая часть шкалы), с нашей точки зрения, указывают на преобладание негатива в восприятии указанных видов сексуального поведения и, следовательно, повышение доли социально желательных ответов. Значения выше указанных критериев (правая часть шкалы) означают положительное отношение к конкретным видам сексуального поведения и довольно высокую степень искренности при ответе респондентов на соответствующие вопросы.

Как свидетельствуют результаты проведенного исследования, влияние метода сбора данных не зафиксировано при ответах по тем трем видам рискованного сексуального поведения, которые большинство респондентов не одобряют: (1) случайные сексуальные связи, (2) сексуальные связи со случайными партнерами при наличии постоянного партнера и (3) сексуальный контакт с партнером своего пола.

А вот оценки тех практик, которые большинством респондентов воспринимаются нейтрально или положительно (средний баллы выше 3,60 балла), обнаружили зависимость от метода опроса (таблица 1); к ним относятся (1) сексуальные контакты в нетрезвом состоянии, (2) наличие большого числа сексуальных партнеров, (3) сексуальные контакты без использования презервативов.

Таблица 1

Распределение оценок рискованных сексуальных практик, не осуждаемых респондентами в зависимости от метода сбора данных, *средние баллы*

Виды анкетирования	Средние баллы отношения		
	к сексуальным контактам в нетрезвом состоянии*	к большому числу сексуальных партнеров	к сексуальным контактам без использования презервативов
Индивидуальное раздаточное	4,55	4,09	5,33
Групповое раздаточное	3,50	4,24	4,70
Групповое компьютерное	2,98	3,02	4,17
По массиву в целом	3,67	3,78	4,73

* Здесь и далее приводятся результаты, имеющие статистическую значимость различий на уровне не менее 95 %.

Во всех трех случаях максимальные оценки дает индивидуальное анкетирование, средние оценки — групповое раздаточное, а самые низкие (самые неискренние) — групповое компьютерное анкетирование. По-видимому, присутствие третьих лиц в последнем случае вынуждает респондентов чаще давать социально желательные ответы, что приводит к занижению оценки.

Можно говорить о тенденции, когда виды сексуального поведения, сопряженного с высокой степенью риска, воспринимаются молодежью как вариант нормы при ответе на вопросы стандартной бумажной анкеты, а при самозаполнении на компьютере расцениваются скорее негативно. Применительно к перечисленным вопросам гипотеза о том, что компьютерное анкетирование повышает искренность респондентов при обсуждении сенситивных вопросов, не подтверждается; скорее верно обратное.

Для оценки уровня риска, допускаемого молодыми людьми в сексуальном поведении, в анкету были включены вопросы о сексуальных практиках самих респондентов. Поэтому влияние метода сбора данных было проанализировано и на этой группе вопросов. В первую очередь, необходимо обратить внимание на вопросы, показавшиеся опрошенным молодым людям особо деликатными. Вопросы, вызвавшие наибольшее смущение, представлены в таблице 2.

Самыми деликатными респондентам показались вопросы об их собственном сексуальном поведении. При анализе влияния метода сбора данных на распределение ответов о наличии сексуального опыта была выявлена статистически значимая зависимость¹. При ответе на вопрос в режиме группового компьютерного анкетирования доля лиц, отрицавших наличие опыта сексуальных отношений, увеличивалась с 7 % (при групповом раздаточном анкетировании) и 13 % (при персональном анкетировании) до 28 %.

Распределение ответов на вопрос об опыте сексуальных контактов со случайными партнерами не зависит от метода сбора данных. Однако гендерная принадлежность здесь имеет значение: мужчины в два раза чаще, чем женщины сообщают о вступлении в сексуальные контакты со случайными партнерами². Эта зависимость повторяется в ответах стандартной бумажной анкеты и группового анкетирования.

Таблица 2

Вопросы, вызвавшие наибольшее смущение респондентов при ответе, %

Вопросы анкеты	
Частота занятий сексом в течение последнего месяца	16,5
Количество сексуальных партнеров за последний год	14,0
Наличие опыта сексуальных отношений	13,2
Допустимо ли вступать в сексуальные отношения с партнером своего пола	8,3
Использование средств контрацепции	7,4
Наличие/отсутствие опыта сексуальных контактов с партнером своего пола	4,1
Отношение к сексуальным контактам с партнером своего пола	2,5
Наличие/отсутствие постоянного сексуального партнера	2,5
Допустимо ли вступать в сексуальные контакты в нетрезвом состоянии	2,5

Распределение ответов о частоте употребления алкоголя перед вступлением в сексуальные отношения зависимости от метода сбора данных не выявило. В среднем этой практики придерживаются 64 % респондентов, что свидетельствует о значительном уровне риска, допускаемом молодыми людьми в своем сексуальном поведении. А вот распределение ответов на вопрос о наличии опыта сексуальных контактов с несколькими партнерами одновременно выявило влияние метода сбора данных³. Так, в режиме персонального анкетирования 19 % указали на наличие такого сексуального опыта, а при групповом анкетировании — лишь 3,5%, однако компьютеру об этом сообщили 57 %, что проинтерпретировать довольно трудно.

Кроме того, выявлена слабая, но значимая взаимосвязь между уровнем материального положения респондентов и опытом вступления в сексуальные отношения с несколькими партнерами одновременно⁴. Среди молодых людей, оценивших свое материальное положение как хорошее, этот показатель составляет 31 %, среди респондентов со средним и ниже материальным положением о наличии опыта группового секса сообщили 16 %.

Еще один показатель — частота измен. Как выяснилось, данный индикатор чувствителен к методу сбора данных. При компьютерном анкетировании доля респондентов, указывавших на

¹ Коэффициент контингенции = 0,241, $p \leq 0,01$

² Коэффициент Крамера V = 1,260, $p \leq 0,001$.

³ Коэффициент Крамера V = 0,508, $p \leq 0,001$.

⁴ Коэффициент Крамера V = 0,162, $p \leq 0,01$.

частую измену сексуальному партнеру, составила 3 %, при индивидуальном раздаточном анкетировании — 17 %¹. Как выяснилось, метод сбора данных не оказал влияния на ответы респондентов о сексуальных практиках с партнером своего пола. В целом о наличии гомосексуального опыта заявили 13 % респондентов, большинство из них девушки (18 %) с хорошим материальным положением³.

Полученные результаты позволяют сделать следующие выводы. Гипотеза о том, что по сравнению с персональным и групповым раздаточным анкетированием групповое компьютерное анкетирование повышает уровень искренности ответов на вопросы о сексуальном поведении, подтвердилась только в двух случаях: при ответе на вопросы о наличии опыта групповых сексуальных практик и о частоте измены постоянному партнеру. Во всех остальных случаях, где было зафиксировано влияние метода сбора данных на ответы респондентов, компьютерное анкетирование снижало качество данных.

Возникает вопрос: какие факторы могли повлиять на снижение уровня искренности респондентов в ситуации опроса на компьютере и привести к заниженным результатам?

К причинам, способствующим снижению искренности респондентов при компьютерном анкетировании, можно отнести присутствие «значимых других», одноклассников, друзей, слишком малое расстояние между компьютерами (вероятно, некоторые респонденты опасались, что их ответы могут увидеть соседи), потенциальная возможность идентификации ответов с личностью респондента, так как опрос проводился в аудиториях, в которых респонденты занимались постоянно. Кроме того, в процессе наблюдения за тем, как воспринимают студенты ситуацию опроса в режиме компьютерного анкетирования, неоднократно было отмечено смущение респондентов, когда требовалось указать специальность, т.е. именно этот вопрос символизировал для них возможность идентификации ответов с личностью. Впоследствии этот вопрос был изменен, и респондентам нужно было указать только курс обучения в вузе.

Таким образом, метод группового компьютерного анкетирования не имеет преимуществ при исследовании рискованного сексуального поведения. В данном случае более качественные результаты обеспечивает метод персонального раздаточного анкетирования. Групповое компьютерное анкетирование, напротив, способствует ухудшению качества данных самоотчетов респондентов о сексуальном поведении.

Однако наряду с описанными недостатками метод обладает рядом достоинств. Например, значительные усилия, которые тратятся на тренировку и инструктаж интервьюеров, можно не прилагать, переложив на компьютер многие человеческие функции и уменьшая шансы на ошибки интервьюеров, а также описки, опечатки, неточности, неразборчивость почерка и т.п. Снимается и ряд вопросов, связанных с ответами не по существу, с ошибками ввода данных, кодировкой. Кроме этого, в зависимости от замыслов исследователя снимаются проблемы незапланированного восприятия анкеты, пропуска ответов, значительно экономятся человеческие, временные, материальные ресурсы. Вполне возможно, что устранение указанных ранее факторов, негативно влияющих на желание респондентов давать искренние ответы по деликатной проблематике, позволит выявить преимущества компьютерного анкетирования по сравнению с другими методами сбора данных.

Измерение базовых жизненных ценностей: сравнение методик и результатов Ценностного исследования Шварца и Европейского социального исследования

Грязнова Ольга Станиславовна, *ИС РАН, НИУ ВШЭ*
Руднев Максим Геннадьевич, *НИУ ВШЭ, ИС РАН*

Один из первых инструментов для измерения жизненных ценностей в массовых опросах в 1950-е годы разработали К. Клакхон и Ф. Стродтбек⁴. Позже было создано большое число ме-

¹ Коэффициент контингенции = 0,259, $p < 0,05$.

² Коэффициент контингенции = 0,140, $p < 0,05$.

³ Коэффициент контингенции = 0,128, $p < 0,05$.

⁴ Kluckhohn F. R., Strodtbeck F. L. Variations in value orientations. Evanston, IL: Row, Peterson, 1961.

тодик и сконструировано множество различных ценностных индексов, фиксирующих ценностные установки в различных социальных группах. На сегодняшний день одним из наиболее авторитетных подходов к сравнительному измерению ценностей является методология Ш. Шварца. На ее основе были разработаны две методики: Ценностный опросник и Портретный опросник, которые использовались в проведении исследований в десятках стран. Первый применялся в Ценностном исследовании Шварца (опросы учителей в 78 странах), а второй — в Европейском социальном исследовании (опросы населения по национальным репрезентативным выборкам в 33 европейских странах). Обе методики использовались для вычисления одних и тех же ценностных индексов. Сравнение результатов двух опросов представляется возможным поскольку, они были проведены в рамках общей методологии (но по разным методикам и разным выборкам) и в ряде случаев в одних и тех же странах. В связи с этим интерес представляет как сравнение самих методик и структур ценностных индексов, так и соотнесение конкретных содержательных результатов.

Цель исследования — обоснование подобия двух методик измерения ценностей Шварца (Ценностного исследования Шварца и ценностного блока Европейского социального исследования) и сходства полученных на их основе результатов.

В теории базовых жизненных ценностей, разработанной Ш. Шварцем, под ценностями понимаются «желаемые, выходящие за рамки конкретных ситуаций цели, отличающиеся друг от друга по значимости и являющиеся руководящими принципами в жизни людей». Именно «конечные цели» выступают критерием группировки ценностей в несколько устойчивых ценностных «доменов», которые в свою очередь организованы в универсальную для всех культур круговую структуру (рис. 1). На основе круговой структуры Шварц выводит два ценностных измерения (оси), которые в максимально обобщенном виде репрезентируют все ценности. Ключевой характеристикой, обеспечивающей возможность использования этой методики в различных странах, является воспроизводимость круговой структуры ценностей во всех культурах («универсальность»)¹. Фонтейн и соавторы продемонстрировали сходство структур ценностей, полученных в различных странах, по результатам Ценностного исследования Шварца², а Билски и соавторы³ показали эти сходства на данных Европейского социального исследования.

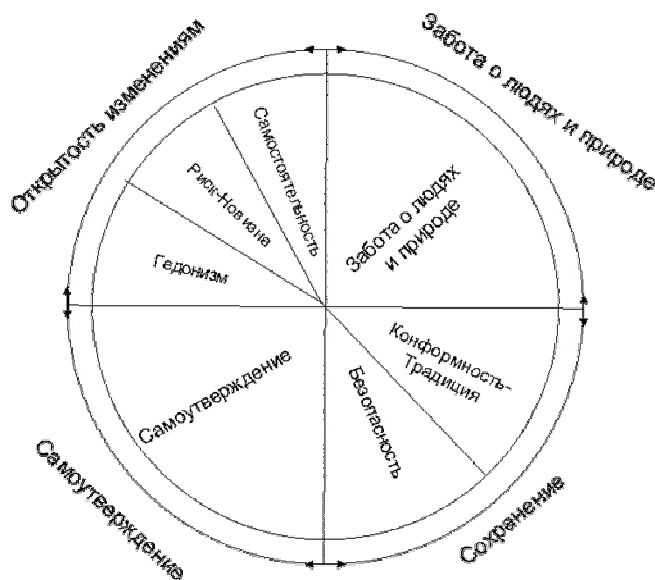


Рис. 1. Теоретические модели взаимоотношений ценностных мотивационных типов

¹ Schwartz S. H. Universals in the content and structure of values: Theory and empirical tests in 20 countries // *Advances in experimental social psychology* / Ed. by M. Zanna. New York: Academic Press, 1992. Vol. 25. P. 1–65.

² Fontaine J.R.J., Poortinga Y.H., Delbeke L., Schwartz S.H. Structural equivalence of the values domain across cultures // *Journal of Cross-Cultural Psychology*, 2008. No. 39. P. 345–365.

³ Bilsky W., Janik M., Schwartz S.H. The Structural Organization of Human Values: Evidence From Three Rounds of the European Social Survey (ESS) // *Journal of Cross-Cultural Psychology*. 2010. July 20.

При разработке инструментария существенное влияние на Шварца оказали исследования ценностей М. Рокича¹. Значительная часть слов-индикаторов в Ценностном опроснике Шварца была заимствована именно из его списка (21 пункт), а остальные из различных международных исследований, а также из текстов по сравнительному религиоведению. Портретный опросник — это более поздняя версия ценностной шкалы, в которой исключены некоторые методические недостатки, имевшие место в Ценностном опроснике.

Два инструмента отличаются не только количеством оцениваемых респондентами пунктов анкеты, но и формой их представления. Ценностный опросник предлагает респонденту два списка ценностей: первый список — существительных, второй — прилагательных, всего 57 слов и словосочетаний, каждое из которых следует оценить по шкале от «Противоречит моим ценностям» (–1) до «Важнейшая в моей жизни» (+7). Инструкция для респондента выглядит так²:

«Отвечая на этот опросник, Вам следует спросить себя: «Какие ценности более важны для меня, как основополагающие принципы моей жизни, а какие менее важны?». Ниже приведены два списка ценностей. Эти ценности взяты из разных культур. В скобках после каждой ценности дано объяснение, которое поможет Вам лучше понять её значение. Ваша задача — указать, в какой степени каждая ценность является основополагающим принципом Вашей жизни. Используйте приведённую ниже шкалу оценок:

Противоречит моим ценностям	Не важна			Важна			Очень важна	Важнейшая в моей жизни
–1	0	1	2	3	4	5	6	7

Оценка каждой позиции по собственной шкале (а не ранжирование, как делал Рокич), позволила исключить фактор взаимозависимости значений предлагаемых характеристик, вызванной спецификой метода. Введение отрицательного пункта шкалы позволило выявить ценности, «от выражения или распространения которых люди стремятся воздерживаться в своем выборе и поведении»³.

Из анализа были исключены те пункты, которые в разных странах демонстрировали отклонения от теоретической модели и нарушали воспроизводимость структуры ценностей. В анализе использовались лишь 45 ценностей, показывавших одинаковые взаимосвязи во всех странах. Однако в опроснике остались все 57 исходных индексов, поскольку изменения методики могли повлиять на измерительные возможности шкалы в целом.

Портретный опросник, включенный в Европейское социальное исследование, был разработан как более компактный и менее чувствительный к культурным различиям инструмент. Его способность измерять ценности была эмпирически подтверждена даже на тех выборках, на которых более ранний Ценностный опросник не давал удовлетворительных результатов. Высокая эффективность Портретного опросника объясняется конкретизацией формулировок и переходом к проективной технике⁴.

Портретный опросник представляет собой набор из 21 ценностного портрета людей, сходство с которыми респонденту предлагается оценить по шкале из 6 пунктов: от «Очень похож на меня» до «Совсем не похож на меня». Респонденту предлагалась следующая инструкция:

«Сейчас я зачитаю Вам краткие описания некоторых людей. Пожалуйста, послушайте каждое описание и скажите мне, насколько каждый из этих людей похож или не похож на Вас? Для ответа используйте эту карточку».

Очень похож на меня	В значительной степени похож на меня	Немного похож на меня	Совсем чуть-чуть похож на меня	Не похож на меня	Совсем не похож на меня
---------------------	--------------------------------------	-----------------------	--------------------------------	------------------	-------------------------

¹ Rokeach M. The Nature of Human Values. New York: Free Press, 1973.

² Опросы в России проводились под руководством А.Г. Левинсона, И.Г. Дубова и Л.М. Смирнова. Перевод опросника осуществлен А.Г. Левинсоном.

³ Schwartz S. Universals in the Structure and Content of Values: Theoretical Advances and Empirical Tests in 20 Countries // Advances in Experimental Social Psychology / M.P. Zanna (ed.). Orlando, FL: Academic, 1992. P. 22.

⁴ Schwartz S., Melech G., Lehmann A., Burgess S., Harris M. Extending the cross-cultural validity of the theory of basic human values with a different method of measurement // Journal of Cross Cultural Psychology. 2001. Vol. 32. P. 538.

Несмотря на описанное методическое различие двух способов измерения, исходные показатели агрегируются в семь теоретически подобных ценностных индексов и две ценностные «оси». Эти категории и являются объектом нашего сравнения.

Для сравнения были отобраны 25 стран, в которых проводилось как в Европейское социальное исследование (39199 респондентов), так и в Ценностное исследование Шварца (4891 респондент). Для расчета ценностных индексов использовалась общая схема агрегирования исходных переменных в семь ценностных индексов и две ценностные «оси»: сначала исходные оценки объединялись в семь ценностных индексов (Безопасность, Конформность–Традиция, Самостоятельность, Риск–Новизна, Гедонизм, Самоутверждение, Забота о людях и природе), затем укрупнялись до четырех ценностных категорий (Сохранение, Открытость изменениям, Самоутверждение, Забота о людях и природе). На основе этих четырех категорий конструировались две ценностные «оси» (Сохранение – Открытость изменениям, Самоутверждение – Забота о людях и природе).

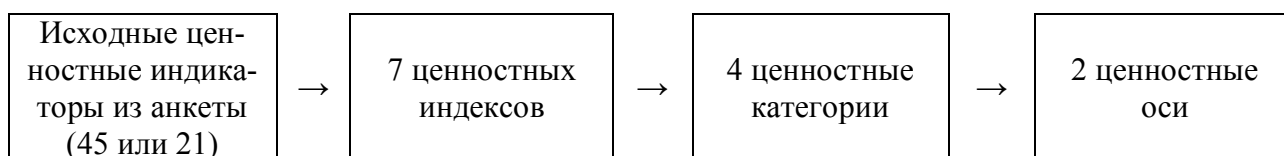


Рис. 2. Последовательность агрегирования показателей.

С помощью **многомерного шкалирования** вслед за Шварцем мы проверили сопоставимость и структуру ценностных индексов. Упомянутые выше Фонтейн и Билски с соавторами помимо сходств структур отмечали и существенные отклонения в различных странах. Кроме того, анализ Ценностного исследования Шварца, проведенный Фонтейном, осуществлялся с нарушениями методологии и с отличиями в технике проверки структуры (другой тип ММШ). Для сравнения структур ценностей в одних и тех же странах, мы провели одинаковые процедуры ММШ (proxsca) внутри каждой из 25 стран и в целом по всем 25 выборкам Ценностного исследования Шварца и Европейского социального исследования. Результаты анализа показали сходство структур в большинстве стран и их соответствие теоретической модели Шварца. В то же время был зафиксирован ряд отдельных отклонений. Однако в целом, надежность и подобие рассматриваемых ценностных индексов была подтверждена. Помимо шкалирования, индексы проверялись на консистентность с помощью коэффициента Альфа Кронбаха. Результаты расчетов продемонстрировали более высокие результаты в Ценностном исследовании Шварца, поскольку для вычисления индексов в нем использовалось большее количество исходных переменных.

Результаты двух исследований мы сравниваем посредством корреляционного и кластерного анализа. Мы рассчитали **корреляции** между средними страновыми значениями 7 ценностных индексов и двух ценностных «осей» по данным Ценностного исследования Шварца и Европейского социального исследования. Они показывают высокую степень взаимосвязи шести из семи ценностных индексов: Безопасность (0,43), Конформность–Традиция (0,48), Самостоятельность (0,45), Гедонизм (0,70), Самоутверждение (0,55) и Забота о людях и природе (0,50). Корреляция между значениями индекса Риска–Новизны является незначимой (0,33). Корреляции между средними значениями по двум осям являются еще более сильными: по оси Открытость изменениям – Сохранение коэффициент корреляции равен 0,71, а по оси Забота о людях и природе – Самоутверждение — 0,54 ($p = 0,01$). Кроме того, были зафиксированы значимые корреляции между разными ценностными индексами, например, Конформность–Традиция из ЦИШ коррелирует с Безопасностью из ЕСИ на уровне 0,73. Однако недиагональные корреляции только на первый взгляд являются отклонением, они наблюдаются только между индексами, входящими в категории более высокого порядка, образующими оси. Результаты корреляционного анализа в целом могут говорить о том, что, несмотря на различие выборок и методик измерения ценностей, эти два независимых международных опроса дают сходные результаты.

Ценностные карты, построенные по данным ЦИШ и ЕСИ по осям Открытость изменениям – Сохранение, Забота о людях и природе – Самоутверждение, подобны друг другу. Наиболее

общий вывод, который можно сделать из визуального анализа карт, заключается в том, что западноевропейские страны находятся на обеих картах в «областях», где доминируют ценности Открытости и Заботы, а постсоциалистические, напротив, — там, где приоритетны ценности Самоутверждения и Сохранения. По результатам дисперсионного анализа, Россия и на карте, построенной по ЦИШ, и на карте ЕСИ находится преимущественно среди постсоциалистических стран.

Кластерный анализ (k-means), проведённый по индивидуальным значениям семи ценностных индексов, показал, что в обоих исследованиях присутствуют сходные ценностные группы. С помощью гар-статистики было установлено, что оптимальное количество кластеров одинаково в обоих массивах и равно четырем. Кластер I составляют респонденты, наиболее приверженные ценностям Открытости изменениям и разделяющие средне-высокую значимость ценностей Заботы о людях и природе. Кластер II объединяет тех, для кого наибольшее значение по сравнению с остальными ценностями имеет Самоутверждение, и чьи ценности занимают срединное положение на оси Открытости–Сохранения. В Кластер III вошли те, чьи показатели по ценностям Заботы являются максимальными в противовес ценностям Самоутверждения, а по оси Открытость изменениям – Сохранение — средними; в Кластер IV — те, для кого характерна особая важность ценности Сохранения и среднее положение по оси Забота о людях и природе – Самоутверждение. Вопреки часто упоминаемой зависимости результатов кластерного анализа от инструмента и особенно от специфики выборки, распределение жителей каждой страны по четырем описанным кластерам оказывается в большинстве случаев схожим по обоим массивам. Например, доли российских респондентов, попавших в различные кластеры по результатам обоих исследований, демонстрируют очень схожую картину: наибольшие группы оказываются в Кластере II и IV, а наименьшие — в Кластерах I и III. Подобные распределения характерны для большинства постсоциалистических стран, в западноевропейских странах большинство опрошенных локализуется в Кластерах I и III, а меньшинство — в II и IV.

Несмотря на различия в методиках и выборках двух исследований, было продемонстрировано сходство, как структур итоговых ценностных индексов, так и их значений в конкретных странах. Полученные результаты легитимируют введение в научный оборот результатов Ценностного исследования Шварца, полученных на нерепрезентативных выборках учителей. Это в свою очередь дает возможность использовать массив данных ЦИШ для анализа ценностей учителей в 78 странах.

Адресная и маршрутная выборки: сравнение результатов

Дюк Елизавета Алексеевна,
Левада-Центр

Для повышения надежности и достоверности информации, получаемой при опросах населения, Левада-Центр совершенствует методики и технологии их проведения. Одним из таких технологических нововведений стал перевод постоянного проекта «Вестник» на адресную выборку, когда на последней ступени отбор респондента производится по случайно выбранным адресам. Эта работа стала возможна благодаря финансовой поддержке Фонда Форда, Фонда Джона Д. и Кэтрин Т. Макартуров и Института «Открытое общество».

Сначала в городах и поселках городского типа, входящих в выборку, были случайным образом отобраны и описаны 128 избирательных участка. (Сельские точки не описывались; опрос в них проходил обычным маршрутным методом.) Затем на описанных участках случайным образом выбраны домохозяйства. Интервьюеру выдавался список адресов домохозяйств и задание на опрос определенного количества респондентов. Внутри домохозяйства респондент выбирался по последнему дню рождения, ближайшему к дате опроса. Интервью должны были проводиться только по выданным спискам; замена адресов не допускалась.

За долгое время проведения массовых опросов в Левада-Центре накоплена статистика недостижимости респондентов по каждому из типов населенных пунктов. Поэтому списки адресов, выдаваемых для опроса, представляют собой задание на опрос определенного количества респондентов по выборке плюс некоторое количество адресов, рассчитанное исходя из средней не-

достижимости респондента по каждому типу. Например, если в каком-то городе нужно было опросить 10 респондентов, а средняя недостижимость по предыдущим опросам составляла 50%, значит, в список для опроса содержал 15 адресов.

Если интервьюер не мог взять необходимое по заданию количество интервью в рамках списка, ему дополнительно выдавалось определенное, очень ограниченное количество адресов. Если он и в этом случае недобирал анкеты, третий раз список не давался. Недобор анкет в одних точках компенсировался перебором в других.

Ранее, до апреля 2010 года, опрос «Вестник» проводился обычным маршрутным методом с фиксацией базовых квартир в соответствии с шагом отбора. Если интервьюер после трехкратного посещения не мог опросить респондента в базовой квартире, он имел право опросить человека в квартире, взятой на замену.

Чтобы оценить различия в характеристиках и мнениях населения, опрошенного по адресной и по маршрутной технологиям, выборка опроса «Вестник-3», проведенного в апреле 2010 года, была поделена на 2 части так, что по маршрутной выборке (вариант 1) в городах и поселках городского типа было опрошено 582 респондента, по адресной выборке (вариант 2) — 571. В итоге мы получили следующие распределения по социально-демографическому блоку.

1. По адресной выборке, по сравнению с маршрутной, было больше опрошено мужчин: 37,3 против 33,2 %.

2. Средний возраст опрошенных одинаковый: 46,4 против 46,3 лет.

3. Структура образования одинаковая.

4. По адресной выборке опрошено меньше специалистов с высшим и средним образованием — 9,8 против 14,0 %.

5. Большой процент опрошенных работает на частных предприятиях — 30,3 против 26,1 %.

6. В среднем «адресное» население на 1000 рублей беднее «маршрутного».

7. По адресной выборке больший процент причисляет себя к средней части среднего слоя (51,0 против 45,8 %) и меньший процент — к низшему слою (9,6 против 15,3 %).

8. По адресной выборке больший процент (45,2 против 40,2 %) считает, что денег хватает на продукты и одежду, но покупка предметов длительного пользования представляет проблему; меньше процент тех, кто едва сводит концы с концами.

9. По адресной выборке в 2 раза больше людей видит себя на 6 и 7 ступеньках общественной лестницы (53 против 27%).

10. Среди «адресных респондентов» больше людей, имеющих подержанный автомобиль (19,4 против 14,9 %).

Также были зафиксированы различия в ответах на содержательные вопросы. В частности, респонденты, опрошенные по методике адресного отбора:

Уровень страны

— чаще выражали неудовлетворенность положением дел в стране и своею собственной жизнью,

— более сдержанно оценили актуальное экономическое положение России и его ближайшие перспективы,

— чаще выражали тревожность и меньшую уверенность в завтрашнем дне,

— более критично оценивали деятельность первых лиц государства,

— более сдержанно выражали доверие первым лицам государства, Государственной Думе, Совету Федерации, органам власти субъектов федерации, политическим партиям, прокуратуре,

— чаще оценивали политическую обстановку в России как напряженную и реже как спокойную,

— реже выражали уверенность в том, что нынешнее правительство России в течение ближайшего года способно улучшить положение в стране,

— реже ожидали улучшения в ближайшие месяцы в политической жизни России,

Уровень семьи

— чаще соглашались с тем, что «приходится “вертеться”», хвататься за любую возможность заработать;

— реже оценивают материальное положение своей семьи как хорошее,

- чаще отмечали, что материальное положение семьи за последний год «скорее ухудшилось»,
- более сдержанно оценивали шансы на улучшение материального положения семьи в будущем,
- чаще выражали ожидания, что цены на основные потребительские товары и услуги в ближайшие один-два месяца будут расти.

Объяснения полученных различий могут быть следующими:

1. Из 128 описанных участков в городах и поселках городского типа примерно 50 расположены на окраинах города с плохой транспортной доступностью. По типу застройки это так называемый «частный сектор»: улицы не освещены, много бараков, домов без коммунальных удобств, непривязанных собак, заброшенных жилищ. Часто это поселки при заводе или при шахте. Проживающее там население социально неблагополучно: много безработных, пожилых людей с низкими доходами. Широкое распространение имеют алкоголизм, наркомания, высокая криминогенность. На некоторых участках наркомания становится проблемой № 1. Ранее эти участки практически никогда не попадали в выборку. Общественный транспорт в эти районы почти не ходит, интервьюеры добираются до домохозяйств с 2–3 пересадками на маршрутных такси или пешком. Кроме того, повышенная криминогенность, бандитизм, пьянство, воровство исключают возможность проводить опрос в наиболее продуктивное вечернее время. Ни один интервьюер не соглашался работать в этих условиях. Зачастую мнение населения, проживающего в таких районах, никогда не изучается. Население там менее «клишировано» в смысле мнений, так как люди проживают в отдельных домах, достаточно отдаленных друг от друга. Они реже общаются с соседями, чем жители обычных микрорайонов с многоэтажной застройкой, где при общении во дворах, магазинах, других общественных местах люди прислушиваются к мнениям друг друга и наиболее конформные из них перенимают мнения большинства. Также в выборку попали несколько районов элитной застройки (коттеджи, пентхаусы, дома повышенной комфортности). Опрос в этих районах тоже почти никогда не проводился.

2. Опрос по маршрутной выборке с возможностью замен недостижимых домохозяйств неизбежно приводит к сдвигам в структуре опрашиваемого населения в сторону лиц нетрудоспособного возраста, домохозяек, инвалидов, женщин, находящихся в отпуске по уходу за ребенком и т.д. Опрос по строго фиксированным адресам позволяет опросить именно того респондента, который попал в выборку. Поэтому в адресной выборке больше мужчин трудоспособного возраста, молодежи и людей с незаконченным средним образованием.

Еще одно достоинство адресной выборки заключается в повышении достижимости респондента по сравнению с маршрутной технологией.

Достижимость по Вестнику, %

	06.2009	08.2009	10.2009	12.2009	02.2010	04.2010	06.2010	08.2010	10.2010
Маршрутная технология	27.6	23	24	23.5					
Адресная технология					29	31	43	43	45

При маршрутном методе интервьюер, сталкиваясь с любой недостижимостью, отмечает ее как отказ и берет адрес на замену. Если и в замененной квартире он сталкивается с недостижимостью, он еще раз ее заменяет, и так до бесконечности. В связи с этим мы имеем очень большую недостижимость по замененным адресам. При адресном методе, в связи с ограниченностью списка адресов интервьюер не злоупотребляет возможностью приписывать респонденту отказ при любом случае недостижимости. Таким образом, *адресная выборка более надежно контролируется.*

К недостаткам адресного метода отнесем следующие:

1. Растягиваются сроки поля: при выборке 1600 респондентов для опроса нужно 2 уик-энда, чтобы полностью пройти 1-й список с 3-х-кратным посещением и еще 1 неделя для опроса по 2-му списку.

2. Непредсказуемость территориальной структуры выборки: в одном и том же населенном пункте в одном опросе выборка может быть сделана полностью, в другом с недобором, в третьем с перебором.

3. Адресная выборка дороже маршрутной. Многократное посещение требует больших транспортных расходов, особенно удорожают опрос многократные поездки в выездные точки. Цена анкеты для интервьюера тоже выше, так как он делает больший объем работы.

«Вестник», 3–5 кратное посещение респондентов, %

Набрано	06.2009	08.2009	10.2009	12.2009	02.2010	04.2010	06.2010	08.2010
при 1-м визите	82	82.7	81.5	83.5	82.6	81	77.3	77.4
при 2-м визите	12.7	11.2	13.8	11.5	12	14	15	16.5
при 3-м визите	5.5	6.1	4.7	5.1	5	4.3	6	7
при 4-м визите					0.2	0.25	1.4	0.9
при 5-м визите					0	0.	0.7	0

Работа с адресной выборкой требует более тщательного подбора и подготовки интервьюеров. Они должны быть готовы к работе с «трудными» интервьюерами на «трудных» участках в вечернее время и в субботу и воскресенье. Очень низкая достижимость респондента в районах с элитной застройкой. Зачастую невозможно попасть не только в подъезд дома, но и во двор с охранниками. Так, в Санкт-Петербурге в выборку попали 3 элитных участка, на которых не удалось опросить ни одного респондента. Тем не менее, еще одно достоинство адресной выборки в том, что она позволяет получать более надежную и достоверную информацию.

Тестирование эффективности выборочных моделей

Журавлева Ирина Валерьевна,
Ивановский государственный энергетический
университет им. В.И. Ленина

Легитимность индуктивных обобщений в отношении объекта на основе выборочных исследований зачастую ставится под сомнение даже профессионалами. И дело тут не в отсутствии доверия выборочным исследованиям: очевидно, что сплошные исследования социальных объектов, за редким исключением, невозможны и невыполнимы в принципе, а грамотная организация выборки существенно экономит ресурсы и обеспечивает получение дефицитной информации об объекте. Однако при поиске наиболее эффективной выборочной модели, позволяющей максимально упростить процедуру отбора и при этом не проиграть в качестве получаемых данных, даже опытный исследователь часто заходит в тупик. Многообразие исследовательских подходов к организации выборочного опроса лишь отражает отсутствие одного наиболее эффективного способа отбора респондентов на все случаи жизни¹.

Для исследования возможностей различных выборочных моделей в марте 2010 г. мы провели опрос среди жителей г. Иваново по проблемам общественной и политической жизни города

¹ См., напр.: Галицкий Е.Б. Методы маркетинговых исследований. М.: Институт Фонда «Общественное мнение», 2004. С.283-290; Ноэль-Нойман Э., Петерсен Т. Все, но не каждый. Введение в методику демоскопии. М.: МГУКИ, 2007. С. 256-259; Рогозин Д.М. Конформная выборка в торговых центрах // Социологический журнал. 2008. № 1. С. 22-48; Чуриков А.В. Случайные и неслучайные выборки в социологических исследованиях // Социальная реальность. 2007. №4. С. 89-109; Яушева А.А. Формирование выборки: поиск новых возможностей // Социологические исследования. 2007. № 8. С. 90-96.

(было опрошено 250 человек)¹. Сбор данных осуществлялся с помощью метода персонального интервью.

Мы сфокусировали исследовательский интерес на смещении опросных данных при использовании разных выборочных моделей. В ходе опроса нами было протестировано пять типов выборки по 50 человек:

- 1) простая случайная выборка (одноступенчатая модель);
- 2) трехступенчатая модель со стратификацией на первой ступени, случайным отбором внутри страт на второй ступени и с применением метода квот на третьей ступени;
- 3) трехступенчатая модель с вышеописанной схемой реализации, только на третьей стадии мы заменили квотный отбор методом Л. Киша;
- 4) квотная выборка;
- 5) конформная выборка.

Для реализации *простого случайного отбора* (n=50) в качестве основы выборки мы использовали электронную паспортную базу данных, в которую занесена информация о жителях г. Иваново. Для случайного отбора респондентов был использован генератор случайных чисел. При составлении выборочной модели на каждого потенциального респондента было дополнительно отобрано по три «запасных». Интервьюеры получили инструкции, согласно которым в случае отсутствия потенциального респондента (либо его пожелания участвовать в опросе чуть позднее) они производили трехкратное посещение место жительства респондента, и только затем обращался к запасному варианту.

Для построения *трехступенчатой выборочной модели* мы первоначально разбили генеральную совокупность на четыре страты, соответствующие четырем административно-территориальным единицам (районам) г. Иваново. Объем каждой страты отражал ее действительную долю в генеральной совокупности. На второй ступени интервьюеры производили случайный отбор домохозяйств внутри каждой страты с использованием маршрутной рандомизации: на каждом из участков случайным образом выбиралась начальная точка маршрута интервьюера (основное условие — не менее двух точек в каждом районе). Последующий выбор дома и квартиры осуществлялся через шаг; в случае состоявшегося интервью полевой работник покидал дом и переходил к следующей единице отбора.

Внутри домохозяйства для отбора конкретного респондента были использованы два разных метода: *метод квот* (n=50) и *карты Л. Киша с двухвопросной техникой отбора*² (n=50).

В качестве квотных параметров для реализации *простой квотной модели* выборки мы использовали пол и возраст респондентов. Опрос проводился в поквартирном режиме, но интервьюеры при отборе людей строго придерживались только квотного задания.

Конформная модель выборки (n=50) была реализована в местах массового скопления людей. Для опроса в городе были выбраны пять точек, отбору подлежали все лица, согласившиеся принять участие в интервью, при этом интервьюеры не делали никаких предпочтений относительно выбора конкретного человека. Для этого каждому интервьюеру было рекомендовано стоять на одном и том же месте, не меняя направления поворота, а по завершении очередного интервью они должны были обратиться к первому же человеку (даже если тот находился в компании), идущему в строго встречном по отношению интервьюеру направлении.

В качестве «внешнего» контрольного эталона для оценки смещений, вызванных разными выборочными стратегиями, мы использовали данные массового опроса (n=1000), проведенного кафедрой социологии ИГЭУ среди жителей г. Иваново по идентичному набору вопросов в тот же период. Для контроля смещений по социально-демографическим показателям нами были использованы данные официальной статистики (Всероссийской переписи населения 2002 г.) и банки данных кафедры социологии ИГЭУ.

¹ Автор выражает благодарность студентке 5 курса специальности 040201«Социология» Г.И. Валунгевичуте за помощь в подготовке полевых работ и обработке результатов и студентам 2–4 курсов за помощь в проведении интервью.

² См.: Галицкий Е.Б. и др. Указ. соч. С. 283.

Сравнение полученных нами данных с результатами массового опроса, проведенного по той же анкете в г. Иваново, позволяет предположить, что наибольшую репрезентативность в целом по блоку содержательных и в особенности социально-демографических вопросов продемонстрировал трехступенчатый отбор людей с реализацией на последней стадии либо схемы Л. Киша, либо квотной модели (табл. 1).

Таблица 1

Доли отклонений выборочных статистик от контрольных данных по всем вопросам анкеты, %

Тип выборочной модели	Отклонения ответов на содержательные вопросы		Отклонения по социально-демографическим показателям	
	Максимальные из наблюдаемых	Минимальные из наблюдаемых	Максимальные из наблюдаемых	Минимальные из наблюдаемых
Простая случайная	16,2*	0,6	16,0*	0,1
Трехступенчатая со стратифицированным отбором (квотный отбор на последней ступени)	18,2*	0,2	10,3	0,0
Трехступенчатая со стратифицированным отбором (схема Л. Киша на последней ступени)	14,0	0,0	9,6	0,0
Квотная	24,4**	0,9	19,6**	0,0
Конформная	31,9**	1,5	30,0**	0,4

* для $\varphi^* p \leq 0,05$.

** для $\varphi^* p \leq 0,01$.

Таблица 2

Общее количество статистически значимых отклонений выборочных статистик от контрольных данных, абс. числа

Тип выборочной модели	Общее количество статистически значимых отклонений	
	По блоку содержательных вопросов	По социально-демографическим показателям
Простая случайная	10	3
Трехступенчатая со стратифицированным отбором (на последней ступени - схема Л. Киша)	4	0
Трехступенчатая со стратифицированным отбором (на последней ступени - квотный отбор)	4	1
Квотная	22	4
Конформная	32	14

Оптимальным вариантом, снижающим вероятность искажения данных об изучаемом объекте, представляется использование трехступенчатой стратифицированной модели выборки с реализацией на заключительной стадии отбора схемы Л. Киша, либо квотного метода.

Детальный дифференцированный анализ распределений ответов на содержательные вопросы выявляет очевидные преимущества «усложненной» трехступенчатой схемы отбора респондентов: именно эта модель (с любой из двух протестированных заключительных ступеней отбора) обеспечивает наименьшие расхождения в распределениях ответов по сравнению с массовым опросом.

Таблица 3

Распределения ответов респондентов на вопрос о партийных предпочтениях
в зависимости от типа выборочной модели, %

Партия, за которую проголосовали бы	Тип выборочной модели										Результаты массового опроса (контрольные данные)
	Простая случайная		Трехступенчатая со стратифицированным отбором (на последней ступени - схема Л.Киша)		Трехступенчатая со стратифицированным отбором (на последней ступени – квотный отбор)		Квотная		Конформная		
	%	Отклонения	%	Отклонения	%	Отклонения	%	Отклонения	%	Отклонения	
Единая Россия	68,0	+12,3	52,0	-3,7	52,0	-3,7	38,0	+17,7	32,0	-23,7**	55,7
КПРФ	14,0	+4,1	8,0	-1,9	12,0	+2,1	22,0	+12,1*	6,0	-3,9	9,9
ЛДПР	8,0	+0,9	2,0	-5,1	10,0	+2,9	16,0	+8,9	28,0	+20,9**	7,1
Справедливая Россия	4,0	-0,7	6,0	+1,3	8,0	+3,3	2,0	-2,7	14,0	+9,3	4,7
Правые силы	0,0	-0,7	0,0	-0,7	2,0	+1,3	16,0	+15,3*	4,0	+3,3	0,7
Против всех	0,0	-6,5	18,0	+11,5	6,0	-0,5	0,0	-6,5	12,0	+5,5	6,5
Не пойду на выборы	0,0	-	4,0	+4,0	6,0	+6,0	6,0	+6,0	2,0	+2,0	0,0
Затруднились ответить	6,0	-1,9	10,0	+2,1	4,0	+3,9	0,0	-7,9	0,0	-7,9	7,9
Пропустили вопрос	0,0	+1,7	0,0	-1,7	0,0	-1,7	0,0	-1,7	2,0	+0,3	1,7

* для ϕ^* $p \leq 0,01$;

** для ϕ^* $p \leq 0,001$.

Применение же конформной модели выборки приводит к наибольшим смещениям по всем блокам вопросов (касающихся и общегородских проблем, и партийно-политических предпочтений, и электорального поведения, и медиапредпочтений и многих других). Так, в табл. 3 представлены данные, убедительно демонстрирующие абсолютную несостоятельность конформной модели выборки при исследовании рейтинга политических партий и движений. Кроме того, судя по данным, представленным в таблице 1, конформная выборка приводит к статистически значимым расхождениям в социально-демографических показателях¹.

Между тем квотная выборка в «чистом» виде и простая случайная выборка, как это ни странно, также приводят к значительным смещениям информации (табл. 1 и 2).

Указанные обстоятельства позволяют усомниться в обоснованности безусловного использования конформных моделей выборки в практике массовых опросов. Подобного рода упрощение организационных форм опроса, конечно, значительно облегчает работу полевого персонала, но при этом актуализирует вопрос об уровне репрезентативности получаемых данных.

Целесообразность использования квотного и простого случайного отбора также может быть поставлена под сомнение. Реализация же простой случайной модели выборки существенно увеличивает временные затраты.

Отказы в телефонных опросах

Журавлева Светлана Львовна,
Ивановский государственный энергетический
университет им. В.И. Ленина

Ошибка, вызываемая неответами — одна из основных составляющих общей ошибки измерения в социологическом исследовании. Высокая доля неответов повышает вероятность систематических смещений в оценке важных для исследователя характеристик изучаемой совокупности. Величина этих смещений зависит и от того, какая доля запланированных респондентов была опрошена, и от того, насколько респонденты и нереспонденты различаются по своим ключевым, значимым с точки зрения исследования, характеристикам. Если эти различия невелики, даже высокий уровень неответов не приводит к возникновению систематических смещений. Напротив, относительно высокий уровень участия в исследовании сам по себе не гарантирует высокой надежности результатов, если отказ от опроса связан с социально-демографическими или иными характеристиками опрашиваемых². Влияние доли отказавшихся на результаты исследования зависит, таким образом, от двух факторов: степени расхождения социально-демографических характеристик отказавшихся от опроса и участвующих в нем, и связи этих характеристик с переменными, подлежащими изучению в исследовании.

Ч. Ибсен и Дж. Баллвэг отмечают, что одним из достоинств использования телефонного интервью как метода сбора опросных данных является высокий уровень участия. По разным оценкам, он варьирует от 85 до 99 % и в целом оказывается существенно выше, чем при использовании персонального или, особенно, почтового опроса³. Однако другие авторы менее

¹ Отклонения выборочных статистик от генеральных параметров оценивались на основании данных Всероссийской переписи населения 2002 г. В качестве независимых контрольных признаков были выбраны также род занятий респондентов, их материальное положение и социально-образовательный статус. За основу в этом случае мы принимали статистику распределения по этим показателям, накопленную в течение многих лет в банках данных кафедры социологии ИГЭУ.

² Goldstein K.M., Jennings M.K. The effect of advance letters on cooperation in a list sample telephone survey // Public Opinion Quarterly. 2002. Vol. 66. № 4. P. 608; Johnson T.P., Cho Y.I., Campbell R.T., Holbrook A. Using community-level correlates to evaluate nonresponse effects in a telephone survey // Public Opinion Quarterly. 2006. Vol. 70. № 5. P. 705; Keeter S., Miller C., Kohut A. et al. Consequences of reducing nonresponse in a national telephone survey // Public Opinion Quarterly. 2000. Vol. 64. № 1. P. 126.

³ Ibsen C.A., Ballweg J.A. Telephone interviews in social research: Some methodological considerations // Quality and Quantity. 1974. Vol. 8. № 1. P. 185.

оптимистичны в оценке метода телефонного интервью, и по их данным, уровень участия не превышает 75 %¹. По оценкам зарубежных специалистов, приводимым Д. Дилманом с соавторами, величина показателя отказов для телефонных опросов варьирует от 9 до 36 %.² Дж. Хокс и Э. де Лиу в результате вторичного анализа данных 45 исследований на разные темы, пришли к выводу, что самый высокий уровень участия достигается в персональном интервью (73 %), за ним следует телефонное интервью (70 %) и почтовый опрос (68 %)³. Более низкий уровень кооперации в телефонном интервью по сравнению с персональным согласуется с представлением о пониженной способности аудиального канала коммуникации передавать эмоции общающихся сторон.

Анализируя динамику уровня отвечаемости, Р. Куртин, С. Прессер и Э. Сингер отмечают, что этот показатель с 1979 г. по 1996 гг. снизился с 72 до 60 %, а к 2003 г. до 48 %⁴. За этот же период среднее число звонков, необходимых для проведения интервью, выросло более чем вдвое — с 4 до 8 %⁵. При этом, как установил Дж. О'Нейл в специальном экспериментальном исследовании, среди неответивших преобладают лица, менее образованные и имеющие более низкий профессиональный статус. Для них характерны более низкие доходы и меньшая социальная активность. Это представители старших возрастных групп (особенно старше 65 лет)⁶.

С этими данными согласуются и выводы Н.А. Романович: отказывающийся от участия в опросах — это человек, принадлежащий к старшим возрастным группам, с низким уровнем образования, как правило, женщина или овдовевший супруг⁷.

Наиболее полный обзор социологических теорий, объясняющих неучастие в опросе, представил Дж. Гойдер. Он рассматривает опрос как инструмент сбора информации в доминантной группе, при этом периферийные, маргинальные группы оказываются исключенными из выборки. Эти группы представлены в большей пропорции среди нетелефонизированных групп населения. Их исключение из выборочной совокупности может привести к достижению более высокого уровня участия в телефонном интервью по сравнению с персональным. Нейтрализует это возможное достоинство телефонного интервью тот факт, что этот метод нарушает приватность жилища, возможно, даже сильнее, чем непосредственный контакт. Дж. Гойдер использует также теорию обмена для описания того, как внимание к нуждам и желаниям респондента помогает повысить уровень участия. Он рассматривает повторные звонки не как повторные вторжения в частную жизнь, а как способ продемонстрировать значимость отобранного индивида для исследования, что может стимулировать дополнительные реакции (как жест взаимности)⁸.

Д. Дилман с соавторами рассматривают в качестве факторов, детерминирующих уровень участия в телефонном опросе, содержание вводной части интервью, использование предварительных писем, а также пол интервьюера⁹. Р. Гроувз, Э. Сингер и А. Корнинг называют

¹ См., например: *Groves R.M., Magilavy L.J.* Increasing response rates to telephone surveys: a door in the face for foot-in-the-door // *Public Opinion Quarterly*. 1981. Vol. 45. № 3. P. 346–358; *Keeter S., Miller C., Kohut A. et al.* Consequences of reducing nonresponse in a national telephone survey // *Public Opinion Quarterly*. 2000. Vol. 64. № 1. P. 125–148; *Singer E., Van Hoewyk J., Maher M.P.* Experiments with incentives in telephone surveys // *Public Opinion Quarterly*. 2000. Vol. 64. № 1. P. 171–188.

² *Dillman D.A., Gallegos J.G., Frey H.* Reducing refusal rates for telephone interviews // *Public Opinion Quarterly*. 1976. Vol. 40. № 1. P. 67.

³ *Hox J.J., De Leeuw E.D.* A comparison of nonresponse in mail, telephone and face-to-face surveys // *Quality and Quantity*. 1994. Vol. 28. № 4. P. 330.

⁴ *Curtin R., Presser S., Singer E.* Changes in telephone survey nonresponse over the past quarter century // *Public Opinion Quarterly*. 2005. Vol. 69. № 1. P. 88.

⁵ *Curtin R., Presser S., Singer E.* P. 90.

⁶ См.: *O'Neil M.J.* Estimating the nonresponse bias due to refusals in telephone surveys // *Public Opinion Quarterly*. 1979. Vol. 43. № 2. P. 228–229.

⁷ *Романович Н.А.* Ситуация опроса глазами респондента // *Социологические исследования*. 1999. № 2. С. 124–125.

⁸ *Groves R.M.* Theories and methods of telephone interview // *Annual Review of Sociology*. 1990. Vol. 16. № 2. P. 226.

⁹ *Dillman D.A., Gallegos J.G., Frey H.* Reducing refusal rates for telephone interviews. P. 68.

также тему исследования и использование вознаграждений¹. М. Тражо, Р. Гроувз, Дж. Лепковски отмечают, что возможности исследователя повысить уровень участия обусловлены темой исследования, особенностями изучаемой совокупности, прилагаемыми усилиями по «конверсии отказов», длительностью полевого этапа исследования, информацией о респондентах и т.д.² Степень влияния каждого из этих факторов в каждом конкретном случае зависит от его важности для данного потенциального респондента, а также от характера его воздействия (позитивного или негативного) на участников исследования³.

Значимость вводной беседы обусловлена тем, что большинство отказов от интервью происходит именно в этот период: около 40 % во время первых нескольких фраз, 50 % после них, но до начала самого интервью, и только около 10 % — после того, как интервью началось⁴. В ходе вступительной беседы опытный интервьюер, озвучивая просьбу об участии в опросе, может наблюдать за поведением собеседника, делать выводы о его специфических потребностях и интересах и приспосабливать под них свои реплики и замечания⁵. Однако ограниченная канальная способность телефона как способа коммуникации лишает интервьюера значительной части информации и затрудняет такое «приспособление». Кроме того, для телефонных опросов с использованием RDD характерен так называемый «cold contact» — контакт без подготовки, когда интервьюер ничего не знает о потенциальном респонденте⁶. Тем не менее, Р. Гроувз свидетельствует, что даже в этих условиях опытные интервьюеры могут достигать более высокого уровня кооперации⁷.

Л. Оксенберг и Ч. Кэннел также подчеркивают роль интервьюера в повышении уровня участия в опросе. Особое значение они придают паралингвистическим характеристикам речи интервьюеров — высоте голоса, его громкости, темпу речи и т.д.⁸

Многие авторы приводят данные о том, что авторитет исследователя способствует увеличению доли согласий в ответ на просьбу об интервью. Авторитет означает поддержку этой просьбы со стороны легитимных властных органов. В опросе авторитет демонстрируется через предварительные письма, идентификационные значки интервьюеров, устную презентацию и т.п. В телефонном опросе при построении выборочной совокупности с применением процедуры RDD доступен лишь последний способ⁹.

Г. Вигдерхаус для минимизации количества отказов рекомендует выбирать оптимальное время дня, день недели, месяц года для проведения телефонного опроса. Результаты проведенного им исследования демонстрируют, что самый высокий уровень участия наблюдается в феврале (64 %), а самый низкий в декабре (41 %). Доля завершенных интервью оказывается наивысшей в период с 18.00 до 18.59 практически во все месяцы года. Время дня играет наибольшую роль в обеспечении высокого уровня завершенных интервью. День недели не имеет большого значения (опросы по выходным не проводились). Наиболее предпочтительным сезоном является весна, затем осень. Лето наименее предпочтительно¹⁰.

¹ Groves R. M., Singer E., Corning A. Leverage-saliency theory of survey participation: Description and an illustration // *Public Opinion Quarterly*. 2000. Vol. 64. № 2. P. 299–308.

² См.: Traugott M.W., Groves R.M., Lepkowski J.M. Using dual frame designs to reduce nonresponse in telephone surveys // *Public Opinion Quarterly*. 1987. Vol. 51. № 4. P. 522–523.

³ См.: Groves R. M., Singer E., Corning A. Leverage-saliency theory of survey participation: Description and an illustration. P. 299.

⁴ См.: Oksenberg L., Cannel C. Effects of Interviewer Vocal Characteristics on Nonresponse // *Telephone Survey Methodology* / Ed. by R.M. Groves, P.P. Biemer, L.E. Lyberg. N.Y.: John Wiley & Sons, 1988. P. 258.

⁵ См.: Groves R. M., Singer E., Corning A. Leverage-saliency theory of survey participation... P. 299.

⁶ Traugott M.W., Groves R.M., Lepkowski J.M. Using dual frame designs to reduce nonresponse in telephone surveys. P. 524.

⁷ Groves R.M. Theories and methods of telephone... P. 232.

⁸ Oksenberg L., Cannel C. Effects of Interviewer Vocal Characteristics on Nonresponse // *Telephone Survey Methodology* / Ed. by R.M. Groves, P.P. Biemer, L.E. Lyberg. N.Y.: John Wiley & Sons, 1988. P. 258.

⁹ Groves R.M. Theories and methods of telephone interview... P. 225.

¹⁰ Vigderhous G. Scheduling phone interviews: A study of seasonal patterns // *Public Opinion Quarterly*. 1981. Vol. 45. № 2. P. 250–259.

Между тем М. Лош с соавторами не обнаружила подобных закономерностей. Только анализ завершённых интервью показал различия в среднем количестве попыток по всем временным периодам (месяц, квартал, время года). Так, летом, особенно в июле, требуется больше звонков для получения завершённого интервью, чем в другие сезоны. Возможно, это связано с тем, что респонденты, склонные к участию в опросах, труднее доступны в определённые периоды времени (лето, третий квартал, июль), тогда как те, кто не склонен участвовать в опросах, одинаково доступны в течение всего года¹. Сезонных различий в уровне кооперации и уровне отказов не было обнаружено. Различия в результатах этих двух исследований могут быть вызваны как факторами методического характера (различия в их дизайне), так и реальными изменениями, произошедшими за 20 лет. Так, повышение занятости женщин и общее повышение активности как женщин, так и мужчин могло привести к снижению доступности респондентов независимо от времени. Это, в свою очередь, могло привести к меньшим различиям в уровне кооперации и отказов по сезонам, кварталам, месяцам, чем это было 20 лет назад².

Рост кооперации под влиянием размера предлагаемого вознаграждения может иметь два объяснения. Во-первых, экономический обмен: оплата рассматривается как компенсация трудностей интервью. Во-вторых, подарок, предлагаемый опросной организацией, стимулирует чувство обязанности ответить тем же (то есть, сотрудничать, отвечать согласием на просьбу об исследовании)³. Эта теория позволяет дать объяснения ранее выявленным закономерностям: респонденты, более заинтересованные в теме исследования, обнаруживают меньший эффект вознаграждения; чем выше воспринимаемая респондентами авторитетность организации, проводящей опрос, тем большее влияние на уровень участия оказывают предварительные письма (например, для академических организаций они более эффективны, чем для коммерческих); респонденты, более обеспокоенные чувствительностью темы, активнее реагируют на использование метода сбора данных, обеспечивающего анонимность; в беседе с респондентами, испытывающими недоверие к организации, проводящей исследование, целесообразно подчеркивать не ее авторитетность или легитимность, а другие особенности исследования; те, кто обеспокоен конфиденциальностью, могут предпочитать самостоятельное заполнение вопросника методом интервью⁴.

Однако далеко не все исследователи однозначно позитивно относятся к использованию каких-либо приемов повышения уровня участия в опросе, полагая, что это повышение может в действительности лишь усугубить проблему смещений⁵. Это может произойти в том случае, если эти приемы повышают уровень участия лишь отдельных групп респондентов. Таким образом, различия между участниками и неучастниками опроса лишь увеличиваются.

Тем не менее, большинство специалистов считает необходимым применение специальных приемов повышения уровня кооперации.

Одним из самых эффективных методов повышения уровня участия в телефонных опросах принято считать предварительные письма⁶. Они предоставляют респонденту информацию о предстоящем звонке, дают ему время на принятие решения и тем самым способствуют снижению негативного воздействия опросной ситуации. Повышение уровня участия в телефонном опросе после получения письма варьирует от 4 %⁷ до 13 %⁸. Однако есть сведения и об

¹ Losch M.E., Maitland A., Lutz G. et al. The effect of time of year of data collection on sample efficiency: an analysis of behavioral risk factor surveillance survey data // *Public Opinion Quarterly*. 2002. Vol. 66. № 4. P. 603.

² Ibid. P. 605.

³ Ibid. P. 303.

⁴ Ibid. P. 307.

⁵ Goldstein K.M., Jennings M.K. The effect of advance letters on cooperation in a list sample telephone survey. P. 609.

⁶ Ibid. P. 608–617; Leeuw E. De, Callegaro M., Hox J., et al. The influence of advance letters on response in telephone surveys: A Meta-analysis // *Public Opinion Quarterly*. 2007. Vol. 71. № 3. P. 413–443.

⁷ Dillman D.A., Gallegos J.G., Frey H. Reducing refusal rates for telephone interviews. P. 66–78.

⁸ Mann C. Do advance letters improve preselection forecast accuracy? // *Public Opinion Quarterly*. 2005. Vol. 69. № 4. P. 561–571; Traugott M.W., Groves R.M., Lepkowski J.M. Using dual frame designs to reduce nonresponse in telephone surveys. P. 522–539.

отсутствии эффекта этого приема. Так, в исследовании Р. Куртина, С. Прессера и Э. Сингера письма, разосланные 44 % респондентов, адреса которых были доступны, не вызвали повышения уровня ответов по выборке в целом, хотя и могли, по мнению авторов, препятствовать его дальнейшему снижению¹.

Имеются также данные о том, что предварительные письма не только повышают уровень участия, но и способствуют получению более демографически точной выборки. В частности, К. Голдстейн и М. Дженнингс, сравнив социально-демографические характеристики респондентов, получивших и не получивших письма, пришли к выводу, что первая подвыборка более точно воспроизводит возраст, пол и партийную принадлежность всех отобранных единиц, чем вторая².

Эффективность писем, по свидетельствам западных ученых, зависит от их длины (они должны быть не слишком длинными), заголовка, времени от получения письма до телефонного контакта (если письмо прочитано задолго до звонка, его эффективность снижается) и от того, кому письмо адресовано (семье в целом или одному из ее членов; во втором случае эффективность выше). Повышает эффективность писем и указание на то, что опрос будет служить интересам значимых социальных групп.

При составлении предварительных писем рекомендуется соблюдать следующие принципы:

1) принцип взаимного обмена, предполагающий заблаговременное предоставление информации в надежде на положительный ответ;

2) принцип связности, консистентности: получатель письма должен воспринимать свое участие в исследовании как связанное с другими обязанностями, например, как часть своих функций избирателя. Для этого в письме обозначается проблема исследования, этой же цели может служить пояснение относительно способов отбора данного респондента³;

3) принцип помощи: в письме необходимо подчеркнуть, какие именно социальные группы получат наибольшую пользу от участия респондента в исследовании.

В качестве альтернативы письмам некоторые исследователи предлагают использовать для предварительного уведомления почтовые карточки. Их использование дешевле, кроме того, они воспринимаются менее официально, чем письма. Однако в карточку можно включить меньше информации, чем в письмо, а времени, которое тратится на ознакомление с ней, может оказаться недостаточно для надежного усвоения представленной информации. В целом по соотношению затрат и результатов почтовые карточки лучше, чем отсутствие каких-либо предварительных уведомлений, но хуже, чем письма.

Таким образом, уровень отказов от участия в телефонных интервью имеет тенденцию к росту, а участники и неучастники исследований различаются по значимым социальным характеристикам.

На уровень участия в телефонных опросах влияют как факторы, непосредственно связанные с опросной ситуацией, так и имеющие социетальную природу.

Достижение сравнительно высокого уровня участия возможно при соблюдении ряда организационных принципов. Привлечение опытных интервьюеров способствует повышению уровня кооперации. В случаях, когда это невозможно, рекомендуется проводить обучение и тренинги полевого персонала, уделяя особое внимание умению «приспособить» вводную беседу к интересам и потребностям конкретного респондента.

Эффективным методом повышения кооперации в телефонных опросах являются предварительные письма. Кроме того, они, по видимому, способствуют приближению социально-демографической структуры выборки к структуре генеральной совокупности.

¹ Curtin R., Presser S., Singer E. Changes in telephone survey nonresponse over the past quarter. P. 95.

² Goldstein K.M., Jennings M.K. The effect of advance... P. 612.

³ Goldstein K.M., Jennings M.K. The effect of advance... P. 610–611.

Опыт использования регрессионного моделирования и EM-алгоритма для заполнения пропусков в данных массового социологического опроса

Зангиева Ирина Казбековна, НИУ ВШЭ

Работа посвящена иллюстрации практического применения и сравнительному анализу эффективности двух алгоритмов импутирования — регрессионного моделирования пропущенных значений и EM-алгоритма (метода максимизации правдоподобия, от англ. Expectation Maximization).

Выбор для рассмотрения именно регрессионного моделирования пропусков и EM-алгоритма обусловлен 3 основными причинами. Во-первых, они широко распространены в зарубежной исследовательской практике, но слабо описаны в отечественной литературе в контексте приложения к пропущенным данным. В российской традиции упор делается на описании их теоретических оснований и использовании для анализа данных с целью получения содержательных выводов. Во-вторых, оба алгоритма достаточно универсальны подходят для реконструкции значений как интервальных, так и порядковых переменными. И, в-третьих, они легко доступны для применения, так реализованы в модуле Missing value analysis (MVA) статистического пакета SPSS.

Работа имеет три основные задачи:

- 1) продемонстрировать применение 2 методов заполнения пропусков: регрессионного моделирования пропущенных значений и EM-алгоритма;
- 2) сравнить эффективность этих методов;
- 3) разработать рекомендации по практическому применению EM-оценивания и регрессионного моделирования как методов заполнения пропусков в данных.

Эмпирическую базу работы составляют данные Национального обследования благосостояния домохозяйств и участия в социальных программах (НОБУС), соответствующие опросу отдельных респондентов (не домохозяйств) осенью 2003 года, представляющие собой репрезентативную всероссийскую выборку. Прежде чем переходить к изложению методики и результатов методического эксперимента, являющегося основой данной работы, следует дать краткое описание сути, достоинств и недостатков использованных для реконструкции пропусков регрессионного моделирования и EM-алгоритма.

EM-алгоритм

EM-алгоритм, основанный на максимизации ожиданий, позволяет не только восстанавливать пропущенные значения с использованием двухэтапного итеративного алгоритма, но и оценивать средние значения, ковариационные и корреляционные матрицы для количественных переменных.

EM-алгоритм в самом общем смысле представляет собой итерационную процедуру, предназначенную для оптимизации некоторого функционала, через аналитический поиск экстремума целевой функции. Этот алгоритм реализуется в 2 этапа, первые буквы названий которых образуют аббревиатуру, обозначающую весь алгоритм:

На *этапе E (expectation)* по совокупности имеющихся абсолютно или частично (по целевой переменной) наблюдений рассчитываются условные, ожидаемые значения целевой переменной для каждого неполного наблюдения. Затем, после получения массива полных наблюдений, оцениваются основные статистические параметры: меры средней тенденции и разброса, показатели взаимной корреляции и ковариации переменных.

На E-этапе по совокупности полных наблюдений определяется функция условного математического ожидания логарифма полной функции правдоподобия при известном значении целевой переменной X вида $Q(\Theta; \Theta^{(m)}) = E_{\Theta^{(m)}}[\log f_{\Theta}(X, Y)|X]$. После определения этой вида этой функции начинается второй этап работы алгоритма.

На *втором этапе M (maximization)* задача алгоритма максимизировать степень взаимного соответствия ожидаемых и реально подставляемых данных, а также соответствия структуры импутированных данных структуре данных для полных наблюдений.

В классическом варианте формально задачу алгоритма по максимизации ожидания можно выразит следующим образом: $\Theta^{(m+1)} = \arg \max_{\Theta} Q(\Theta; \Theta^{(m)})$. Здесь Θ обозначает рассчитанное ожидаемое условное значение, для отсутствующей характеристики некоторого наблюдения.

Количество итераций, которые совершает алгоритм, прежде чем выдать окончательный результат, максимизирующий ожидания, определяется формальным критерием остановки, в котором однако присутствует субъективный элемент, поскольку исследователь сам заранее выбирает некоторую метрику $\rho(\Theta_1, \Theta_2)$ и положительное число ε , стремящееся к бесконечно малому.

Итерация алгоритма с порядковым номером m становится последней, если значение метрики, показывающей приращение в условных ожидаемых значений, не превышает заранее заданного значения ε . Формально, условие остановки EM-алгоритма записывается как: $\rho(\Theta^{(m)}, \Theta^{(m-1)}) < \varepsilon$

Это означает, что приращение после итерации (m) является не значимым и бесконечно малым, поэтому следует остановиться на результате итерации ($m-1$) и считать их окончательными¹.

EM-алгоритм импутирования обладает следующими достоинствами:

- 1) дисперсия значений переменной сохраняется на уровне максимально близком к исходному;
- 2) применим для импутирования всех видов переменных;
- 3) позволяют оценивать не только пропущенные значения, но и средние значения, ковариационные и корреляционные матрицы для количественных переменных.

К его недостаткам можно отнести то, что:

- 1) метод относительно субъективен и требует значительного вмешательства исследователя в выборе аналитической модели;
- 2) алгоритм итерационный, и с каждой итерацией увеличивается длительность процесса импутирования и объем обрабатываемых данных.

Регрессионное моделирование пропусков

Один из способов использования алгоритмов регрессионного анализа для заполнения пропусков предложил У. Бака. Его метод заключается в подстановке средних, условных по присутствующим в наблюдениям.

Если переменные Y_1, \dots, Y_k (у которых есть пропуски в данных) распределены по многомерному нормальному закону со средним μ и ковариационной матрицей Σ , то регрессия пропущенных значений для определенного, рассматриваемого объекта линейна по присутствующим значениям с коэффициентами, которые являются известными функциями от μ и Σ . Бака предлагает сначала оценивать μ и Σ по полным наблюдениям, а затем, используя полученные оценки, строить линейные регрессионные модели для предсказания пропущенных значений по присутствующим для каждого наблюдения.

Подставляя значения переменных, присутствующих для данного наблюдения, в регрессионное уравнение, получаем прогноз пропущенных переменных для этого наблюдения. Пропуски, восстановленные с использованием метода Бака, обеспечивают адекватные оценки средних значений, при условии справедливости предположение о нормальности наблюдений. Полученная выборочная ковариационная матрица, рассчитанная по восстановленным методом Бака данным, занижает величину дисперсии и ковариаций. Однако это завышение не столь значительно, как в случае замены пропусков на обычные безусловные средние значения².

В большинстве случаев импутирование при помощи регрессионных моделей осуществляется в два этапа. На первом этапе на совокупности полных наблюдений отстраивается рег-

¹ Королев В.Ю. EM-алгоритм, его модификации и их применение к задаче разделення смесей вероятностных распределений: Теоретический обзор. М.: ИПИ РАН, 2007. С. 12–13.

² Злоба Е., Яцкив И. Статистические методы восстановления пропущенных данных // Computer Modeling & New Technologies. 2002. Vol. 6. No. 1. P. 53.

рессионная модель, и оцениваются коэффициенты в уравнении, где в качестве зависимой переменной выступает целевая переменная, чьи пропущенные значения необходимо восстановить. Затем в полученное уравнение подставляются известные значения независимых переменных предикторов, и для каждой целевой переменной рассчитывается отсутствующее значение. В случае интервальных и абсолютных переменных рассчитывается конкретное значение, а для порядковых и номинальных переменных с некоторой вероятностью предсказывается категория, к которой должен быть отнесен объект.

Классическое регрессионное моделирование пропущенных значений имеет узкую область применения, которая формально ограничена только переменными, измеряемыми на уровне не ниже интервального. Однако в социологических исследованиях количественные переменные в чистом виде используется редко. Чаще переменные измеряются по порядковым или номинальным шкалам, что на первый взгляд делает регрессионное моделирование не применимым в данной ситуации. Но ограничение на уровень измерения переменных для регрессионного моделирования можно преодолеть за счет создания для порядковых и номинальных переменных фиктивных переменных, которые уже отвечают всем требованиям к данным. Существует несколько модификаций классической регрессионной модели, которые можно использовать для заполнения пропусков по переменным, измеренным по шкалам ниже интервального уровня¹.

К достоинствам регрессионного моделирования, применяемого к реконструкции пропущенных значений можно отнести:

1) универсальность алгоритма: есть регрессионные модели, подходящие для всех типов переменных; регрессионное моделирование применимо для реконструкции и количественных, и категориальных переменных;

2) простоту расчетов и интерпретации результатов;

А в качестве недостатков следует упомянуть:

1) чувствительность результатов импутирования к наличию в данных артефактов;

2) зависимость качества импутирования от выбора переменных-предикторов; при слабой объяснительной силе предикторов точность импутирования получается низкой. Для правильного выбора предикторов необходимо еще до обучения модели проанализировать все возможные предикторы и выбрать те, которые сильнее всего коррелируют с целевой (зависимой) переменной;

3) для каждого целевого объекта должны быть известны значения всех предикторов, в противном случае импутирование не возможно.

Содержание и результаты методического эксперимента

Методика эксперимента

Поставленные задачи были решены в рамках методического эксперимента, реализованного по следующей схеме:

1. Из массива данных Национального обследования благосостояния домохозяйств и участия в социальных программах (НОБУС, 2003 год) были отобраны все наблюдения, которые содержали полные (без пропусков) данные по характеристикам, включенным в логистическую регрессионную модель для предсказания вероятности быть вовлеченным в сектор неформальной занятости в зависимости от социально-демографических характеристик.

2. На полученной подвыборке полных наблюдений были оценены параметры распределений изучаемых переменных и построена искомая логистическая регрессионная модель. Таким образом, был получен идеальный случай — эталон для сопоставления результатов после импутирования.

3. В имеющийся массив полных наблюдений были искусственно внесены полностью случайные пропуски (наблюдения в которые были внесены пропуски отбирались с помощью генератора случайных чисел) по социально демографическим переменным в 4-х вариантах:

¹ Крыштановский А.О. Анализ социологических данных с помощью пакета SPSS. М.: ГУ-ВШЭ, 2006. С. 164–165.

количество пропусков по каждой переменной, идентичное исходному массиву НОБУСа, а так же 20, 40 и 60 % пропусков каждой переменной. В итоге получилось 4 массива данных разной степени полноты.

4. На каждом из полученных 4 неполных массивов была построена искомая логистическая регрессионная модель и оценены смещения в оценках, возникшие в результате разного количества пропусков в данных. Сразу отметим, что потери были неизбежны, так как логистический регрессионный анализ предполагает автоматическое исключение всех неполных наблюдений, однако было необходимо установить их реальные масштабы.

5. Затем пропуски, имевшиеся в каждом из 4 массивов были заполнены при помощи 2 методов: регрессионного моделирования и EM-алгоритма.

6. Для оценки эффективности импутирования каждого метода были вновь оценены параметры распределений изучаемых переменных и количество правильных предсказаний (точных подстановок) каждым и снова построена искомая регрессионная модель.

7. Результаты, полученные при использовании каждого из методов заполнения пропусков, были сопоставлены с идеальными значениями полных наблюдений. Более эффективным объявлялся тот алгоритм импутирования, результаты которого были ближе к идеальным.

Основные результаты эксперимента: потери информации

На подготовительном этапе с помощью процедуры MissingValueAnalysis пакета SPSS все переменные исходного массива были проанализированы на предмет масштабов, структуры пропусков и возможных схем сопряженности (совместных пропусков) между пропусками в отдельных переменных. Это было сделано для того, чтобы повторить эту структуру на искусственных пропусках, которые будут внесены в массив.

По результатам MVA оказалось, что пропуски в исходном массиве не взаимосвязаны и случайны. Были рассчитаны доли пропусков по каждой переменной в исходном массиве (далее «Репрезентативные пропуски»).

На каждом из 5 массивов данных — без пропусков и с каждым из 4 наборов пропусков — была построена уже описанная модель влияния социально-демографических характеристик человека на его вероятность быть занятым неформально. В каждом случае определялась доля объектов, которые исключались из анализа из-за наличия у них хотя бы одного пропуска по включенным в модель переменным.

Таблица 1

Изменение качества модели при использовании только полных наблюдений в логистическом регрессионном анализе

Переменная	Доля правильных предсказаний в модели, %				
	Данные без пропусков	Подвыборки с репрезентативными пропусками			
		5%	20%	40%	60%
Формальная занятость	97	97	96	96	96
Неформальная занятость	48	47	56	60	72
Общее качество модели	92	91	91	92	93
Из анализа исключено из-за пропусков, % объектов	0	23	68	92	99

Пороговое значение вероятности для логистической модели = 0,5.

Всего объектов в массиве полных наблюдений: 41447.

На первый взгляд кажется, что с ростом количества пропусков в данных ситуация только улучшается и качество модели повышается, так как при наличии 60 % пропусков по каждой переменной процент верно предсказанных случаев неформальной занятости вырос на 24 % по сравнению с данными без пропусков. А процент правильного предсказания формальной занятости с ростом количества пропусков практически не меняется и остается на очень высоком уровне 96–97 %.

Однако последняя строка таблицы, в которой приведена доля исключенных в каждом случае из анализа объектов, показывает, что при 60 % пропусков в каждой переменной, т.е. когда качество предсказания максимально, репрезентативность модели сводится почти к нулю, так как из анализа исключается 99 % объектов. Ценность содержательных выводов для исследователя неформальной занятости стремится к 0.

Теперь очевидна необходимость хотя бы с некоторым приближением восстановить недостающую информацию, чтобы уменьшить масштабы исключения объектов из анализа; даже если модели, построенные по восстановленным данным, будут менее точными по сравнению с идеальной моделью. Потери в точности будут зависеть от качества предсказания отсутствующих значений, которое можно будет оценить, так как истинные значения по всем характеристикам для каждого объекта известны. Но, потеряв в точности, можно выиграть в репрезентативности результатов. Предсказательное качество модели и ее репрезентативность, как видно из таблицы 1, находятся в обратной зависимости. Но каждая из этих характеристик одинаково важна для интерпретации результатов, и исследователю необходимо определить для себя их оптимальное сочетание, чтобы оценить необходимость импутирования данных.

При реконструкции пропусков мы использовали разные модификации классической регрессионной модели с фиктивными переменными в зависимости от типа шкалы и размерности целевой переменной:

Зависимые переменные (значения которых восстанавливаются)	Тип регрессионной модели (с фиктивными переменными)
специфический стаж и заработная плата	классическая линейная регрессия
форма собственности фирмы и категория должности	множественная логистическая регрессия
наличие дополнительной работы	бинарная логистическая регрессия

В качестве независимых переменных (предикторов) в виде фиктивных переменных использовались пол и уровень образования индивида, и напрямую — его возраст. Теперь анализируем потери в точности (по сравнению с идеальной моделью — без пропусков), которые возникают при заполнении всех пропусков с использованием разных моделей.

Импутирование интервальных переменных

Данные таблицы 2 показывают, что при минимальной доле искусственно рассчитанных данных значения среднего специфического стажа и средней заработной платы отклоняются от истинного на 0,01 года и 8,5 рублей соответственно.

Таблица 2

Качество импутирования значений интервальных переменных
в зависимости от количества исходных пропусков в данных

		Специфический стаж		Зарплата	
		Среднее	Дисперсия	Среднее	Дисперсия
Данные без пропусков		5,96	14,42	3821,2	8369635
EM-алгоритм	5%	5,95	13,48	3812,7	7628057
	20%	5,99	12,43	3710,3	6724276
	40%	6,13	9,30	3682,0	5212286
	60%	6,18	7,22	3754,7	3828585
Импутирование при помощи регрессионных уравнений	5%	6,29	14,10	3812,9	7726731
	20%	5,26	13,40	3612,2	6757558
	40%	4,76	10,94	3532,4	5287328
	60%	4,09	8,34	3457,1	3898290

С ростом доли искусственных значений смещение от значений среднего при регрессионном моделировании значения переменных с большей дисперсией (здесь, доход) выше, чем у EM-алгоритма. А для переменных с незначительной дисперсией (в нашем случае, специфический трудовой стаж) однозначного вывода сделать нельзя, так как изменения среднего значения для двух методов разнонаправлены: EM-алгоритм приводит к завышению среднего значения стажа, а регрессионное моделирование — к его занижению.

Однако и EM-алгоритм, и регрессионное моделирование даже при минимальном количестве заполняемых пропусков уменьшают дисперсию значений переменных. Если при EM-импутировании 5 % пропусков дисперсия снижается в среднем лишь на 8 %, то при заполнении 20 % пропусков — на 17 %, при 40 % пропусков — 37 % дисперсии, при 60 % пропусков — 52% дисперсии. Регрессионное моделирование снижает дисперсию на 5, 13, 30 и 48 %, соответственно.

Поэтому, если перед исследователем стоит задача оценить степень неоднородности обследуемой совокупности по некоторой интервальной переменной, например по доходу, заполнение пропусков может привести к существенному занижению оценок, по сравнению с ситуацией включения в анализ только полных наблюдений.

Импутирование категориальных переменных

Очевидно, что в случае импутирования значений категориальных (порядковых и номинальных) переменных регрессионное моделирование (множественная и бинарная логистическая регрессия) более эффективна, как с точки зрения большего количества верных подстановок, так и меньших, по сравнению с EM-алгоритмом, потерь дисперсии.

Таблица 3

Качество импутирования значений порядковых и номинальных в зависимости от количества импутируемых значений, %

Переменная	Доля верных предсказаний							
	EM-алгоритм				Регрессионное моделирование			
	5%	20%	40%	60%	5%	20%	40%	60%
Форма собственности фирмы (номинальная)	35	25	14	5	55	46	35	35
Категория должности (порядковая)	20	14	11	13	48	33	25	19
Наличие дополнительной работы (номинальная)	97	97	97	95	98	97	95	95
Изменение дисперсии значений (по сравнению с идеальным случаем)	-3	-16	-35	-52	+1	-5	-26	-40

При регрессионном моделировании теряется на 10 % меньше дисперсии, а прирост в количестве верных подстановок еще более значителен. В зависимости от количества импутируемых значений (от 5 до 60 %) при использовании EM-алгоритма доля правильно предсказанных значений переменной «форма собственности фирмы» снижается с 35 до 5 %, «должности» — с 20 до 13 %, а в регрессионном моделировании с 55 до 35 % и с 48 до 19 %, соответственно.

Изменение качества логистической регрессионной модели

После импутирования EM-алгоритмом и регрессионным моделированием во всех рассмотренных случаях получаются практически одинаковые показатели качества предсказания типа занятости. Расхождения в них для двух методов импутирования не превышают 4 %. Так, регрессионная модель 5 % пропусков в данных предсказывает неформальную занятость на 4% лучше, чем EM-импутирование; при заполнении 20 % пропусков на 3% лучше результат дает EM-алгоритм.

Заключение

В качестве заключения сформулируем рекомендации, основанные на результатах проведенного методического эксперимента по использованию EM-алгоритма и регрессионного моделирования для реконструкции пропусков в данных.

По параметрам изменения средних значений, количества верных подстановок, потерь дисперсии и качества логистических моделей между долей исходных пропусков и качеством импутирования наблюдается обратная взаимосвязь: с ростом количества пропусков точность импутирования снижается. Этот вывод вполне логичен, если принять во внимание, что при увеличении пропусков в данных уменьшается совокупность полных наблюдений, которая используется в качестве данных для прогнозирования отсутствующих значений. Получается, что необходимо предсказывать больше пропусков, используя для этого меньше данных, и сохранить при этом прежний уровень точности оценок не возможно.

В случае количественных переменных, характеризующихся большой дисперсией значений и при увеличении доли пропусков в данных EM-алгоритм является более эффективным методом импутирования, чем регрессионное моделирование, так как дает меньшее отклонение от известных значений, по сравнению с регрессионным моделированием.

В случае однородной по некоторой количественной характеристике совокупности складывается неопределенная ситуация. Здесь изменения мер средней тенденции, рассчитанных на импутированных разными методами данных разнонаправлены: регрессионное заполнение пропусков приводит к занижению среднего значения, а EM-алгоритм — к его переоценке.

Что касается показателей разброса в данных, то оба метода даже при минимальном количестве исходных пропусков занижают дисперсию значений переменных.

Если нужно восстановить значения категориальной переменной (порядковой или номинальной), регрессионное моделирование предпочтительнее для сохранения исходной дисперсии: в среднем теряется на 10 % меньше дисперсии, чем при EM-импутировании, а прирост верных подстановок оказывается даже выше.

Если необходимо оценить степень однородности исследуемой совокупности и пропусков в данных немного, следует ограничиться анализом полных наблюдений, так как импутирование чревато существенным занижением оценок.

Многоуровневый ковариационный анализ эффекта этничности в социальных сетях школьников

Иванюшина Валерия Александровна,
Александров Даниил Александрович,
СПб. филиал НИУ ВШЭ

Настоящая работа является частью долгосрочного проекта Научно-учебной лаборатории социологии образования и науки проекту «Дети-мигранты в российских школах»¹. Этот проект, начатый в 2007 году, сочетает качественные и количественные методы: массовые опросы школьников увязаны с детальным исследованием отдельных школ (по методу «кейс-стади») и глубинными интервью с учениками, учителями, администраторами и родителями. Географически наше исследование охватывает Московскую область и Санкт-Петербург. Эти регионы были выбраны потому, что они являются основными центрами притяжения трудовых мигрантов (как внутренних, так и трансграничных) в Российскую Федерацию. Согласно данным Госкомстата, из всех въезжающих в РФ трансграничных мигрантов в Московскую область прибывает 26 %, в Санкт-Петербург — 12 %. Таким образом, именно в этих городах можно ожидать обнаружить школы с большим количеством детей из семей мигрантов.

В России, в отличие от ряда стран Западной Европы, рост числа детей из семей мигрантов в школах еще только начинается, однако уже сейчас в урбанизированных районах РФ есть школы, где доля иноэтничных мигрантов достигает 40 %. Изучение детей из семей мигран-

¹ Александров Д.А., Баранова В.В., Иванюшина В.А. Дети из семей мигрантов в школах Санкт-Петербурга: предварительные данные. СПб.: Изд-во СПб. Политехнического Университета, 2010.

тов актуально и в прикладном отношении — для принятия решений в системе образования, и в теоретическом — сейчас есть возможность поймать момент становления новых сообществ, который в европейских странах был упущен исследователями несколько десятков лет назад.

Одним из фокусов нашего исследования является изучение адаптации и интеграции иноэтничных детей в российской школе. Традиционно для измерения адаптации и интеграции мигрантов использовались их школьные успехи и образовательные притязания, а также анализ аттитудов — как самих мигрантов, так и принимающего сообщества. В настоящее время, как в исследованиях школы, так и в миграционных исследованиях наряду с классическими статистическими методами широко применяются современные методы анализа социальных сетей¹. Чтобы оценить позицию мигрантов на основе сравнения большого числа сетей, необходим многоуровневый анализ.

Планирование исследований по этой теме затруднено отсутствием базовой статистики о количестве детей из семей мигрантов в российских школах. Официально такая статистика не собирается, а инициативно собираемые данные не отличаются надежностью и сопоставимостью, поскольку, в первую очередь, нет ясности в том, кого следует считать мигрантами и по каким показателям собирать такую статистику.

Предлагаемая статья основана на обработке данных опроса школьников восьмых–десятых классов, проведенного весной 2010 г. в Санкт-Петербурге по стратифицированной выборке. Генеральной совокупностью в нашем исследовании являются государственные общеобразовательные школы Санкт-Петербурга (598 школ). Частные школы, коррекционные школы, начальные школы, интернаты, кадетские корпуса были исключены из анализа.

При конструировании выборки перед нами стояла двоякая задача: получить репрезентативные данные и найти классы с достаточно высокой долей мигрантов, чтобы изучить сетевые эффекты в полиэтничных классах.

Выборка строилась в два этапа. На первом этапе из генеральной совокупности случайным образом были отобраны 30 «обычных школ» и 10 «школ повышенного статуса». На втором этапе была осуществлена довыборка малочисленных школ, т.к. из наших предварительных исследований мы знали, что именно в этих школах концентрируются дети из семей иноэтничных мигрантов. Всего в выборку попали 104 школы из всех 18 районов Санкт-Петербурга. Было опрошено 419 классов (8–10 классы) и собрано 7300 анкет.

Для опросов школьников нами разработан оригинальный инструмент, позволяющий оценивать отношение ученика к школе и образованию, его ощущение включенности в школьную жизнь. В анкету также включены вопросы на сетевые отношения в классе, позволяющие измерять социальную ин/экслюзию непосредственно, а не только путем анализа мнений. Важно подчеркнуть, что опросы проводятся во всем классе целиком, что позволяет осуществлять сравнение между детьми разной этничности и разной миграционной истории.

При разработке анкеты были использованы статьи американских, канадских, британских исследователей, а также опросные инструменты масштабных международных кросс-секционных и лонгитюдных исследований образования и миграции: National Education Longitudinal Study (NELS), США, National Longitudinal Survey of Youth (NLSY), США, Youth in Transition Survey (YITS), Канада, Youth Cohort Study (YCS), Великобритания, Children of Immigrants Longitudinal Study (CILS), США, Programme for International Student Assessment (PISA).

Анкета включала подробный социально-демографический блок, вопросы о материальном положении семьи; вопросы о культурном капитале; вопросы о текущих школьных успехах ученика; об образовательных планах; о выборе будущей профессии. Блок «Миграционная и языковая история» включал вопросы о времени переезда семьи в Петербург; о месте рожде-

¹ Luken V., Tranmer M. Personal support networks of immigrants to Spain: A multilevel analysis // *Social Networks*. 2010. Vol. 32. P. 253–262; Vermeij L. et al. Ethnic segregation in context: Social discrimination among native Dutch pupils and their ethnic minority classmates // *Social Networks*. 2009. Vol. 31. P. 230–239; Baerveldt C. et al. Ethnic boundaries and personal choice: Assessing the influence of individual inclinations to choose intra-ethnic relationships on pupils' networks // *Social Networks*. 2004. Vol. 26. P. 55–74.

ния родителей; о родном языке матери, отца, и самого ученика; о том, на каком языке говорят дома. Вопросы о родном языке помогают устанавливать этничность респондента в тех случаях, когда он не отвечает на вопрос о национальности. В вопросе о времени переезда семьи в Петербург (Подмосковье) предлагались варианты ответа, разбитые на периоды. Это важно, т.к. длительность проживания в принимающей стране является важным фактором, наряду с этнической принадлежностью, влияющим на адаптацию.

Для оценки социально-экономического статуса семьи в анкету были включены открытые вопросы о трудовой занятости отца и матери школьника. Требовалось не только назвать профессию, но и кратко описать, чем они занимаются на своей работе. Этот инструмент позволяет достаточно точно переводить ответы школьников о работе родителей в код ISCO-88 (International Standard Classification of Occupations), который далее перекодировался в шкалу социально-экономического статуса ISEI (International Socio-Economic Index).

Для оценки про/антишкольных настроений оценки вовлеченности учеников в школьную жизнь (sense of belonging) и самооценки популярности в классе в анкету было включено несколько блоков так называемых «аттitudных вопросов», которые оценивались по 4-балльной шкале.

В анкету также входили «сетевые» вопросы для изучения межличностных и межгрупповых отношений в классе. Вопрос звучал следующим образом: «Назови, с кем ты больше всего общаешься в классе»; «Назови, с кем ты меньше всего общаешься в классе». В том и в другом случае можно было выбрать до 10 человек.

Обработка сетевых данных требует специальных методов анализа, отличных от классических статистических процедур, поскольку многие предположения, на которых основана классическая статистика, неприменимы к сетевым данным, в частности, предпосылка о независимости наблюдений (так называемая проблема сетевой автокорреляции). Сетевые данные по определению взаимозависимы. В настоящее время разработано несколько подходов для решения проблемы сетевой автокорреляции. Одним из таких подходов является метод, при котором путем компьютерного моделирования генерируется множество сетей, сохраняющих заданные характеристики исходной сети, и подсчитывается вероятность случайного возникновения того или иного изучаемого эффекта.

Такие алгоритмы стали возможными при вычислительных возможностях современных компьютеров; отметим, что расчет не самой сложной сетевой модели для 150 акторов занимает несколько часов, а с увеличением числа акторов или сетей время вычислений возрастает многократно.

Для изучения сегрегации и социального исключения мы применяли два подхода. Во-первых, использовали социально-психологическую перспективу через изучение мнений и аттitudов: детям задавали вопросы на «чувство принадлежности», на самооценку популярности. В соответствии с ожиданиями, в ситуации успешной интеграции, то есть при отсутствии социальной эксклюзии, самоощущения у детей из семей иноэтничных мигрантов не будут отличаться от таковых этнического большинства.

Во-вторых, мы использовали сетевой анализ, то есть изучали реальные связи и отношения между школьниками в классе. Этот метод позволяет непосредственно измерить интеграцию тех или иных детей в социальную ткань класса и школы.

Для обработки сетевых данных была применена многоуровневая модель p2, реализованная в компьютерном пакете StOCNET. Модель p2 предложили в 2004 г. ван Дуийн, Снайдерс и Зильстрой; в 2006 году модель усовершенствована для многоуровневого анализа, в результате чего стал возможен одновременный анализ нескольких сетей¹. Этот метод позволяет моделировать случайные эффекты на диадных данных, причем в качестве независимых переменных могут выступать как атрибуты акторов, так и свойства диад (диадные ковариаты).

¹ Van Duijn M.A.J. et al. P2: a random effects model with covariates for directed graphs // *Statistica Neerlandica*. 2004. Vol. 58. P. 234–254; Zijlstra B.J.H. et al. The multilevel p2 model: A random effects model for the analysis of multiple social networks // *Methodology*. 2006. Vol. 2. P. 42–47.

Для построения моделей были использованы классы с тремя и более учениками, принадлежащими к этническим меньшинствам. Таких классов в нашей выборке было 80, они распределялись по 53 школам, общее количество школьников в этих классах равнялось 1575. Основные характеристики классов приведены в Таблице.

Подготовка данных для сетевого анализа является отдельным трудоемким этапом работы. Во-первых, необходимо унифицировать все имена, названные респондентами, и сопоставить их с соответствующими атрибутами (характеристиками акторов). Во-вторых, необходимо преобразовать данные из анкет в виде матриц сопряженности в формате, пригодном для ввода в программу StOCNET. Для выполнения обоих этапов мы воспользовались специально написанными скриптами на языке Python.

Таблица

Класс	Min	Max	Mean	Median
Размер	8	29	20,5	20
Число minority	3	10	4,2	4,0
% minority	4,0	43,8	21,3	18,6

В нашем анализе зависимой переменной являлись дружеские связи между учениками. Моделировались вероятности образования связей между школьниками в зависимости от их пола, этничности, социально-экономического статуса семьи, отношения к школе и образованию и т.д. Анализ проводился в несколько этапов: вначале каждая из объясняющих переменных была проверена индивидуально, затем в модель последовательно добавлялись переменные. В конечной модели были сохранены все переменные, для которых был обнаружен статистически значимый эффект.

Интерпретация коэффициентов модели r^2 как коэффициентов логистической регрессии позволила сделать следующие выводы.

- Самый сильный эффект на образование связей оказывает гендерная гомофилия, причем среди мальчиков он выражен сильнее, чем среди девочек. Это известная закономерность, которая неизменно выявляется во всех исследованиях подросткового общения и дружб. Подчеркнем, что все остальные эффекты считались с учетом этого основного влияния.
- Русскоязычные дети (majority) не имеют предпочтений относительно этнической принадлежности своих друзей; этнические меньшинства (minority) проявляют тенденцию образовывать связи с minority.
- С учетом эффектов гендера и этничности была также выявлена гомофилия по анти-школьным аттитюдам и по намерениям относительно получения высшего образования, а именно: дружба более вероятна между учениками с близким уровнем анти-школьной культуры, а также со сходными планами относительно высшего образования.
- Гомофилии по социально-экономическому статусу, чувству принадлежности, намерениям оставить школу после 9 класса не обнаружено.
- Ни одна из групповых характеристик (тип школы, размер школы, число и процент иноэтничных детей в классе) не оказалась значимой в нашей модели, т.е. выявленные закономерности образования дружеских связей не связаны с типом и составом школы.

Изучение отчислений среди студентов бакалавриата/специалитета НИУ ВШЭ

Колотова Елена Васильевна, *НИУ ВШЭ*

Введение

Отчисления студентов, в наиболее широком понимании включающие в себя все случаи выбывания из университета до момента получения диплома, являются популярной темой для исследований как в США, так и в большинстве стран Европы. В этих странах уровень отсева является одним из показателей конкурентоспособности вуза: чем ниже уровень отчислений,

тем более престижным является вуз. Для государства отсев студентов представляет потерю средств (в случае финансирования обучения) и издержки, связанные со снижением производительности отчисленного в будущем. Для самих студентов отчисление из университета представляет собой потерю вложенных в свое образование материальных и временных средств, а также сопутствующие этому психологические трудности.

Уровень студенческих отчислений различается в зависимости от страны. По данным Организации экономического сотрудничества и развития, в России на 2008 год он составлял 21 %. По данным волны 2008 года Российского мониторинга экономического положения, 23 % россиян в возрасте до 25 лет учились в университетах, академиях или вузах, но не получили диплом. Среди жителей всех возрастов, доля имеющих неполное высшее образование, по данным РМЭЗ 2008 волны исследования, составляет 15 %. По данным ОЭСР, в 2008 г. наиболее высокая доля студентов, не закончивших высшее образование, фиксировалась в США (54 %), наименьшая в Дании (16 %) и Корее (15 %)¹.

Методология исследования

Отчисление студента — событие, качественное изменение, происходящее в определенный момент времени. Поэтому для изучения причин и масштабов отчислений необходим особый тип данных — лонгитюдные записи, в которых содержится информация о том, в какой момент событие наступило для индивида или группы индивидов.

Наиболее подходящей методологией для анализа данных о регистрации события является анализ наступления событий (АНС) (event history analysis), позволяющий работать с изменяющимися во времени независимыми переменными и пропусками в данных². Задачей методов АНС является изучение того, как меняется риск наступления события в зависимости от продолжительности нахождения в группе риска³, а также от других значимых характеристик индивида. Это достигается посредством оценки функции выживаемости (survival function) $(S(t))$ — вероятности того, что событие не наступило до момента t ⁴.

Операционализация понятия «отчисление студента»

Объектом исследования является отчисление студентов из университетов. Предметом изучения являются масштаб, закономерности и факторы, влияющие на отчисление студентов. Фокус нашего исследования направлен на изучение масштаба и причин отчисления по причине академической неуспеваемости, поскольку это является одной из основных причин выбытия студентов из рассматриваемого университета. Согласно правилам обучения в Высшей школе экономики, наличие трех задолженностей ведет к автоматическому отчислению студента. В административных данных ВШЭ о студентах (АСАВ) фиксируется несколько типов отчислений студентов, однако две трети случаев приходится на отчисление по академической неуспеваемости.

Заметим, что другие типы отчислений (по собственному желанию, перевод в другой вуз и т.д.) могут в части случаев лишь формально обозначать другую причину, но фактически являться отчислением по академической неуспеваемости. Перевод на другой факультет или уход в академический отпуск также нередко используются студентами как способ избежать отчисления из университета. Поэтому при наличии у студента несданных дисциплин, либо при его последующем отчислении по академической неуспеваемости эти две категории мы также относили к отчислению по академической неуспеваемости.

Итак, один из показателей обозначал выбытие студентов из университета или с факультета по причине академической неуспеваемости. В него вошли как формально зафиксированные случаи отчисления по академической неуспеваемости, так и другие типы отчислений,

¹ Education at a Glance 2010: OECD Indicators. www.oecd.org/edu/eag2010

² Применение традиционных методов анализа к таким данным приводит к потере информации и смещениям в оценках, см.: Allison P.D. Event history analysis. Beverly Hills: Sage, 1984. P. 10–11.

³ В группу риска входят индивиды, которые имели возможность испытать наступление события в течение конкретного временного интервала.

⁴ Handbook of data analysis / Ed. by M. Hardy, A. Bryman. Thousand Oaks: Sage, 2004. P. 371.

перевод на другой факультет или уход в академический отпуск при наличии академических задолженностей у студента (не было итоговой оценки хотя бы по одному предмету¹). Этот показатель в дальнейшем мы будем называть *отчислением по причине академической задолженности*.

В анализе также использовался суммарный показатель «*все переходы студентов*», который включает в себя все типы отчислений, а также уход в академический отпуск, перевод на другой факультет, вне зависимости от того, есть ли у студента академические задолженности или нет. Этот показатель позволяет рассматривать отчисление студентов в более широком смысле: как любое (временное или постоянное) прекращение обучения студентом либо во ВШЭ, либо на том же самом факультете.

В исследовании были поставлены следующие задачи:

- 1) оценить масштаб отсева студентов в НИУ ВШЭ, поступивших в бакалавриат или специалитет в 2007, 2008 и 2009 гг. (по обоим показателям);
- 2) определить время (курс и модуль обучения), когда вероятность студента НИУ ВШЭ быть отчисленным по причине академической задолженности наиболее высока;
- 3) определить, какие характеристики студентов, известные вузу (НИУ ВШЭ), повышают вероятность студента быть отчисленным по причине академической задолженности.

В качестве источника для анализа данных использовалась база статистических данных НИУ ВШЭ о студентах на этапе их поступления и обучения АСАВ — «Абитуриент. Студент. Аспирант. Выпускник». На базе данных были выделены когорты студентов, поступивших на бакалавриат (специалитет) НИУ ВШЭ в конкретном году. При отнесении к когорте рассматривалось только первое зачисление студента в НИУ ВШЭ; если студент отчислялся несколько раз за время своего обучения, рассматривалось только первое отчисление, после которого этот студент из анализа исключался.

При анализе учитывался не только факт отчисления студента из университета (или его перехода, ухода в академический отпуск), но и время наступления данного события. Показателями времени является количество модулей², прошедших с начала зачисления студента в университет до момента наступления события. Отметим, что в анализе дату события мы определяли по записи его статуса об отчислении, что могло отличаться от реальной даты отчисления от нескольких дней до нескольких месяцев. Время между зачислением студента и его отчислением (по причине академической задолженности) корректировалось, если в этот промежуток студент уходил в академический отпуск: его длительность вычиталась.

При анализе уровня отчислений студентов по причине академической задолженности остальные случаи выбытия студентов из университета (по собственному желанию, невыхода из академической неуспеваемости и т.д.) рассматривались как цензурированные, или неполные, данные. Они включались в анализ, но рассматривались до момента прекращения наблюдения за студентами (их ухода из университета). Таким образом, эти наблюдения учитывались при расчете вероятности наступления события (отчисления) в периоды обучения соответствующих студентов в Школе, но не учитывались при расчете рисков отчисления в периоды после их выбытия из университета.

Масштаб и причины отсева студентов в НИУ ВШЭ

В таблице 1 представлена статистика отчислений студентов, как по показателю отчислений по причине академических задолженностей, так и по показателю, учитывающему все переходы студентов. Отметим, что мы не можем сравнивать показатели разных когорт в силу разного периода наблюдения. Также имеющихся данных недостаточно, чтобы установить, какая доля студентов «не доходит» до выпуска.

Однако данные показывают, что 22 % поступивших в 2007 г. студентов по итогам 3 лет обучения были вынуждены покинуть вуз (факультет) *по причине академических задолжен-*

¹ Рассматривалось наличие несданных дисциплин, первая попытка сдачи которых приходилась на период от 6 месяцев до момента наступления события (выбытия или перехода) до дня наступления события.

² Модуль — единица периода обучения в ВШЭ. Для анализа считалось, что 1 модуль равняется 2 месяцам.

ностей; 5 % студентов были отчислены без академических задолженностей (что мы рассматриваем как добровольный уход из университета); еще 5 % студентов без академических задолженностей переходили на другой факультет или уходили в академический отпуск. В целом, доля делавших хотя бы какие-то переходы к началу 2 модуля 4-го курса составила 31 %.

Таблица 1

Уровень отчислений студентов в НИУ ВШЭ для когорт 2007–2009 гг.

Показатели на конец периода наблюдения	Когорта (год поступления в НИУ ВШЭ)		
	2007	2008	2009
	Конец периода наблюдения		
	1 модуль 4 курса	1 модуль 3 курса	1 модуль 2 курса
Всего студентов в когорте в момент поступления	2038	1965	2353
Все переходы: суммарный показатель	31,4	29,1	21,3
Из них:			
1. Выбыли при наличии академических задолженностей	22,0	19,0	12,5
<i>В том числе:</i>			
— отчислены по причине академической неуспеваемости	13,1	10,5	8,0
— отчислены по причине невыхода из академического отпуска, либо по собственному желанию, либо по причине неисполнения условий договора, либо переведены в другой вуз	3,0	3,1	1,4
— переведены на другой факультет, либо предоставлен академический отпуск, либо перешли на ИУП (при наличии академических задолженностей, либо впоследствии были отчислены по академической неуспеваемости)	5,8	5,4	3,1
2. Выбыли без академической задолженности	4,7	5,5	4,7
<i>В том числе:</i>			
— отчислены по причине невыхода из академического отпуска, либо по собственному желанию, либо по причине неисполнения условий договора, либо переведены в другой вуз	4,6	5,4	4,7
— отчислены в связи со смертью	0,1	0,1	0,0
3. Переходили на другой факультет или уходили в академический отпуск, не имея задолженностей	4,7	4,6	4,1

Важный вопрос о том, как изменяется показатель уровня отчислений с течением времени, мы вынуждены оставить за рамками этой статьи, в силу отсутствия подходящих для анализа данных об обучении студентов за 4 года.

На рис. 1 отображено распределение функции риска *быть отчисленным по причине академической задолженности* на протяжении периода обучения студента. Риск в данном случае — это отношение вероятности быть отчисленным к вероятности не быть отчисленным. Диаграмма иллюстрирует, что риски повышаются в периоды пересдач и аккумулируют незачеты, полученные в предыдущих модулях. Наибольший риск быть отчисленным по причине академических задолженностей приходится на начало 2 курса обучения — период пересдачи незачетов, полученных во второй половине первого курса обучения. Таким образом, можно сделать вывод о том, что именно студенты первого курса обучения в наибольшей степени подвержены риску отчисления.

Что увеличивает риск отчисления студентов?

Для выявления факторов, влияющих на вероятность отчисления по академическим задолженностям на первом курсе, использовались данные по студентам когорт 2007–2009 гг. поступления. В качестве независимых характеристик первоначально выступали:

- 1) пол студента;
- 2) место проживания до поступления во ВШЭ (Москва, Московская обл., другой регион);
- 3) тип среднего учебного заведения, которое закончил студент (лицей, гимназия, центр образования, средняя общеобразовательная школа, центр профессиональной подготовки);
- 4) наличие золотой медали за успехи обучения в школе (да/нет);

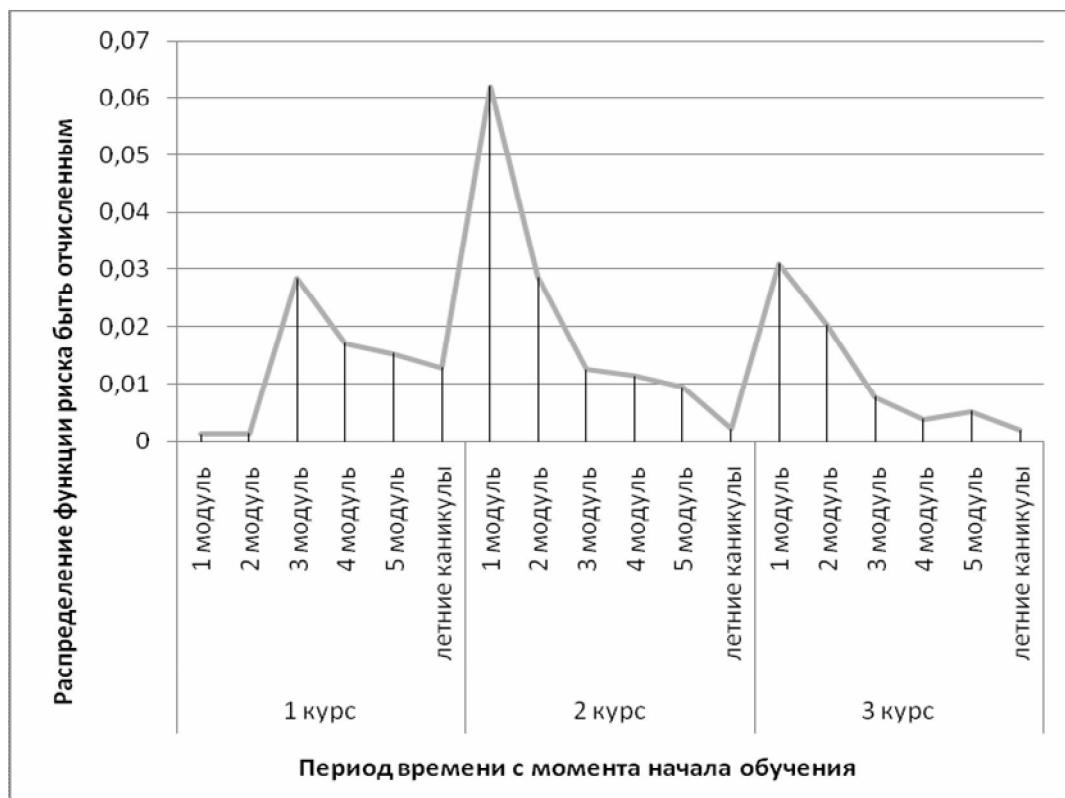


Рис. 1. Распределение функции риска быть отчисленным по причине академической задолженности (суммированы данные когорт 2007, 2008 и 2009 гг.)

- 5) возраст студента на момент зачисления;
- 6) место зачисления студента (бюджетное или коммерческое);
- 7) год зачисления студента во ВШЭ;
- 8) нормированный¹ проходной балл на факультет/отделение, куда был зачислен студент.

В ходе анализа между некоторыми независимыми переменными обнаружили взаимодействия. Поскольку «тип школы» оказался связанным с «местом проживания до поступления во ВШЭ», а «проходной балл» — с наличием золотой медали за школьные успехи и местом, на которое был зачислен студент, из регрессионной модели были исключены переменные «тип школы» и «нормированный проходной балл». Кроме того, поскольку место, на которое студент был зачислен, оказалось связанным с регионом проживания, была сформирована новая переменная с пересечением категорий двух исходных переменных.

Оценочные значения коэффициентов при остальных регрессорах, оказавшихся значимыми, представлены в таблице 2. Интерпретация оценочных значений коэффициентов схожа с интерпретацией нестандартизованных коэффициентов регрессии. Экспонента от коэффициента B для интервальных переменных отражает, на сколько процентов изменится функция риска при увеличении независимой переменной на одну единицу. Так, каждый дополнительный год для возраста студента на этапе его поступления в университет увеличивает вероятность его отчисления на 15 %. Интерпретация значений коэффициентов при категориальных и дихотомических переменных требует учитывать базовую категорию — ту, с которой сравниваются все остальные. Для студентов, которые закончили школу с золотой медалью, риск отчисления составляет 36 % от риска тех, кто закончил школу без золотой медали.

Итак, результаты анализа говорят о том, что студентов, отчисленных и не отчисленных по академической неуспеваемости в течение 3 лет и 2 месяцев обучения, больше всего дифференцируют следующие факторы.

¹ Проходной балл рассчитывался для всех факультетов по 100-балльной шкале.

Оценочные значения коэффициентов при регрессорах
для моделей длительности (регрессия Кокса)

Характеристика студента	Для всех когорт (N=6332)		Для когорты 2007 г. (N=2036)		Для когорты 2008 г. (N=1962)		Для когорты 2009 г. (N=2352)		
	Sig.	Exp(B)	Sig.	Exp(B)	Sig.	Exp(B)	Sig.	Exp(B)	
Возраст (кол-во лет)	0,00	1,15	0,00	1,22	0,02	1,18	0,05	0,84	
Закончили школу с золотой медалью (базовая категория — «без золотой медали»)	0,00	0,36	0,00	0,41	0,00	0,36	0,00	0,32	
Мужской пол (базовая категория «женский пол»)	0,00	1,62	0,00	1,69	0,00	1,57	0,00	1,44	
Поступил на бюджетное отделение и проживает	в Москве (базовая категория)	0,00		0,00		0,00		0,00	
	в другом регионе	0,01	1,38	0,69	0,93	0,01	1,64	0,01	1,98
	в Московской области	0,15	1,27	0,80	0,93	0,77	1,09	0,01	2,20
Поступил на коммерческое отделение и проживает	в Москве	0,00	2,45	0,00	1,99	0,00	2,53	0,00	3,03
	в другом регионе	0,00	2,75	0,00	2,39	0,00	2,40	0,00	4,09
	в Московской области	0,00	2,90	0,00	2,88	0,00	2,40	0,00	3,65

Наличие золотой медали за школьные успехи снижает вероятность отчисления почти в 3 раза. Влияние данного фактора одинаково как для модели, включающей все исследуемые когорты, так и для моделей, построенных для каждой из когорт.

Чем **старше** студент при поступлении в университет, тем выше вероятность его отчисления по причине академической задолженности (учитывалось только первое поступление студента, не рассматривались ситуации поступления в последующие разы, что позволило избежать смещений при оценке данного параметра). Этот вывод справедлив для когорты 2007 г. поступления. При контроле переменной «пол» в когорте 2009 г. зависимости рисков отчисления студентов от возраста не наблюдается. В когорте 2008 г. возраст значим только для представителей мужского пола: каждый дополнительный год повышает риск отчисления в 1,2 раза.

Студенты **мужского пола** имеют примерно в полтора раза более высокий риск отчисления, чем студенты женского пола. Пол студента является одним из важнейших факторов, влияющих на вероятность быть отчисленным, по результатам зарубежных исследований¹: в большинстве стран студенты мужского пола выбывают чаще, чем студенты женского пола. Например, у американских юношей более высокие риски выбытия приходятся на первые курсы обучения, а для девушек — на последние (по семейным обстоятельствам: выход замуж, рождение детей и т.д.). К сожалению, имеющиеся пока данные не позволяют проверить эту гипотезу для студентов ВШЭ.

В когортах 2008 и 2009 гг. среди поступивших на бюджетное отделение, у студентов, **проживающих за пределами Москвы** риски отчисления выше, чем у москвичей. Зачисление студента в университет на **место с оплатой обучения** увеличивает риск его отчисления по сравнению со студентом, поступившим на бюджетное место, для всех когорт студентов.

Влияние результатов ЕГЭ на риск отчисления

Одним из актуальных вопросов высшего российского образования является изучение того, насколько хорошо экзамен по ЕГЭ дифференцирует школьников по их уровню знаний, способностям к обучению, а также то, насколько эффективным является использование ре-

¹ Например: Bowen W.G., Chingos M.M., McPherson M.S. Crossing the Finish Line: Completing College at America's Public Universities. Princeton: Princeton University Press, 2009.

зультатов ЕГЭ в качестве вступительных экзаменов в вуз. Также у руководства университета нет однозначного мнения о том, что является более эффективной формой отбора студентов — олимпиады или прием по результатам ЕГЭ.

Одним из показателей эффективности отбора с помощью того или иного способа может рассматриваться уровень отчислений студентов по академической неуспеваемости. Мы проанализировали, есть ли различия в уровне отчислений:

— в когорте 2007 г. между зачисленными по результатам олимпиады и зачисленными по результатам сдачи экзаменов,

— в когорте 2009 г. между зачисленными по результатам олимпиады и зачисленными по результатам ЕГЭ¹.

И дескриптивный анализ, и оценка функции выживаемости показывают, что различий в уровне и риске отчисления по академической неуспеваемости между данными группами не наблюдается. Различия же наблюдаются между зачисленными по результатам вступительных испытаний (экзамены, ЕГЭ, олимпиады), с одной стороны, и зачисленными по льготам, целевому набору и на коммерческое отделение, с другой. Для последних риск отчисления возрастает в разы (см. таблицы 3 и 4).

Таблица 3

Уровень отчислений студентов по академической неуспеваемости среди студентов когорт 2007 и 2009 гг. в зависимости от способа поступления

Способ поступления	Когорта 2007 г.		Когорта 2009 г.	
	Доля отчисленных, %	Кол-во зачисленных на 1 курс	Доля отчисленных, %	Кол-во зачисленных на 1 курс
Победитель олимпиады	9,3	367	5,8	600
Набрал проходной балл	10,9	312	6,9	534
На основании договора	31,7	1002	18,0	1148
По льготам или целевому набору	32,3	93	21,1	71
Всего	21,9	2037	12,5	2353

Далее на основе данных о студентах когорт 2008 и 2009 года поступления мы проанализировали, насколько хорошо результаты ЕГЭ предсказывают вероятность отчисления студента, дифференцируют неотчисленных и отчисленных по причине академической неуспеваемости. Из анализа мы исключили «стоцентные» результаты по ЕГЭ, полученные благодаря каким-то привилегиям (например, призеру олимпиады), а также результаты тех студентов, которые поступили по результатам олимпиад. Ввиду разницы в проходном балле данные о студентах, поступивших на бюджетное и коммерческое отделение, анализировались отдельно.

Таблица 4

Оценочные значения коэффициентов при предикторе «способ поступления» для моделей длительности (регрессия Кокса)

Способ поступления	Для когорты 2007 г. (N=2036)		Для когорты 2009 г. (N=2352)	
	Sig.	Exp(B)	Sig.	Exp(B)
По проходному баллу на бюджетное место (базовая категория)	0,00		0,00	
По льготам или целевому набору	0,00	3,63	0,00	3,14
Поступил на бюджетное место, используя медаль	0,19	0,53	нет такого способа	
Победитель олимпиады	0,48	0,84	0,44	0,84
На основании договора	0,00	3,33	0,00	2,70
Набрал полупроходной балл и прошел собеседование	0,18	1,42	нет такого способа	

¹ Когорта 2008 г. исключена из анализа из-за сложностей в работе с административной базой данных.

Сравнение средних результатов для когорты 2009 года (таблица 5) показывает, что значимые различия между группой отчисленных и неотчисленных студентов наблюдаются по всем показателям только для студентов коммерческой формы обучения. Для студентов бюджетного отделения значимые различия наблюдаются только по результатам ЕГЭ по иностранному языку.

Таблица 5

Сравнение средних результатов ЕГЭ среди неотчисленных и отчисленных по причине академической задолженности студентов, поступивших во ВШЭ в 2009 году

Показатель		Неотчисленные		Отчисленные		Уровень значимости
		Средний %	Кол-во наблюдений	Средний %	Кол-во наблюдений	
Балл ЕГЭ по русскому языку	Поступил на бюджетное отд.	80,7	622	79,3	54	0,28
	Поступил на коммерч. отд.	73,1	912	69,7	201	0,00
Балл ЕГЭ по математике	Поступил на бюджетное отд.	73,5	378	70,4	37	0,07
	Поступил на коммерч. отд.	66,2	671	60,3	161	0,00
Балл ЕГЭ по иностранному языку	Поступил на бюджетное отд.	86,8	550	78,0	41	0,00
	Поступил на коммерч. отд.	78,8	881	72,1	188	0,00
Балл ЕГЭ по обществознанию	Поступил на бюджетное отд.	76,6	459	75,7	35	0,49
	Поступил на коммерч. отд.	69,7	632	65,8	132	0,00

Таблица 6

Оценочные значения коэффициентов при предикторах «балл ЕГЭ» для моделей длительности (регрессия Кокса)

Характеристика студента		Для когорты 2008 г.			Для когорты 2009 г.		
		Sig.	Exp(B)	Всего наблюдений	Sig.	Exp(B)	Всего наблюдений
Балл ЕГЭ по русскому языку	Поступил на бюджетное отд.	0,00	0,94	661	0,34	0,98	669
	Поступил на коммерч. отд.	0,00	0,96	755	0,00	0,95	1007
Балл ЕГЭ по математике	Поступил на бюджетное отд.	0,02	0,97	490	0,13	0,98	411
	Поступил на коммерч. отд.	0,01	0,98	581	0,00	0,97	730
Балл ЕГЭ по иностранному языку	Поступил на бюджетное отд.	недостаточно наблюдений			0,00	0,97	584
	Поступил на коммерч. отд.				0,00	0,98	964
Балл ЕГЭ по обществознанию	Поступил на бюджетное отд.				0,68	0,99	490
	Поступил на коммерч. отд.				0,00	0,94	706

То же самое относится и к баллам ЕГЭ в качестве предикторов выживаемости для когорты 2009 года (таблица 6). Однако для когорты 2008 года, когда прием по результатам ЕГЭ еще не был массовым, результаты ЕГЭ были значимыми предикторами риска отчисления для студентов бюджетной формы обучения. Можно предположить, что результаты ЕГЭ в 2009 году стали менее показательными для предсказания будущей академической успешности студентов, чем для более ранних когорт студентов.

Заключение

Проведенный анализ показал, что студенты первого курса обучения оказываются в ситуации наибольшего риска отчисления, что подтверждается и другими исследованиями¹. Вероятно, руководству вуза следует создать дополнительные меры адаптации первокурсников к обучению в университете, например, консультации старшекурсников, специальные занятия-консультации преподавателей, персональные тьюторы и т.д.

Выделяются факторы, которые оказывают значимое влияние на риск отчисления для всех когорт студентов, и их оценочные значения практически не меняются с течением времени: пол, регион проживания, место зачисления, способ поступления, наличие золотой медали за школьные успехи. Другие факторы для одних когорт являются значимыми, а для других нет: возраст, результаты ЕГЭ для студентов бюджетного отделения.

Оценивая эффективность разных способов приема студентов с точки зрения риска их отчисления, мы не выявили различий между студентами, поступившими по результатам ЕГЭ, и студентами, поступившими по результатам олимпиады. В дальнейшем планируется определить эффект разных типов олимпиад, а также определить эффективность того или иного способа поступления для разных факультетов.

В когорте 2009 года поступления результаты ЕГЭ являются значимым предиктором риска отчисления студентов только для коммерческих студентов. В то же время, для бюджетного отделения 2008 года поступления результаты ЕГЭ по русскому языку и математике хорошо дифференцируют студентов по показателю отчислений.

Изменения в системе приема студентов в вуз в 2009 году, безусловно, создали новые условия для селекции, однако дать предварительную оценку этим изменениям можно будет лишь после анализа данных по когорте 2010 года.

Ограничением настоящей работы является то, что были протестированы, в основном, внешние факторы, информация по которым была доступна на базе статистических данных университета. В исследование не были включены факторы, которые, согласно существующим моделям отсева студентов, гораздо более тесно связаны с процессом отчисления из университета. В модель необходимо включить вовлеченность в учебный процесс, во внеучебную деятельность, степень социальной интеграции, мотивы обучения в университете, его материальное положение, получение финансовой помощи от других родственников, наличие семейных обязанностей, время, отводимое на работу. Мы надеемся, что проведение специальных обследований студентов поможет их включить в дальнейшее исследование отсева студентов НИУ ВШЭ.

Свободное программное обеспечение в социологических исследованиях

Матрунич Александр Владимирович,
«Матрунич Консалтинг»

Подавляющее большинство программного обеспечения (далее ПО), используемого социологами в исследовательских проектах, относятся к категории закрытых коммерческих продуктов. Текстовые документы верстаются в редакторе Word от компании Microsoft, статистический анализ производится в SPSS от компании IBM. Подобные программы достаточно дорогостоящие, что требует выделения существенного бюджета на их приобретение. Использование нелегальных версий повышает риски исследователя, связанные с законом; в будущем, когда Россия вступит в ВТО, доступ к краденным версиям программного обеспечения станет более затруднительным и опасным². Широкая распространенность пиратских копий коммерческих закрытых продуктов способствует тому, что вне внимания российского исследователя остаётся целый пласт альтернативного программного обеспечения.

¹ Чудиновских О.С., Донец Е.В. Возможности и ограничения завершения высшего образования в элитном вузе: на примере МГУ им. М.В. Ломоносова. М.: МАКС Пресс, 2004.

² Кречетников А. Россию могут снова не пустить в ВТО // Русская служба BBC. 2010. июля 22 // http://www.bbc.co.uk/russian/business/2010/07/100722_russia_usa_wto_rights.shtml

Понятие свободного программного обеспечения

Рассмотрим вехи зарождения свободного программного обеспечения¹ (далее СПО). В период с 50-х до 70-х годов прошлого века для пользователей немногочисленных компьютеров обладание свободами в отношении ПО было нормальной практикой. Производители компьютерного оборудования предоставляли ПО в качестве бесплатного бонуса к основному товару. Пользователи свободно запускали программы, изучали их исходный код, изменили его и передавали модифицированные версии своим коллегам. Компьютерная программа была подобна научной статье, которая была доступна всем желающим, идеи из этой статьи можно использовать в других статьях и т.д.

В конце 1960-х гг. в индустрии информационных технологий в связи с ростом специализации стало появляться всё больше компаний, которые занимались только разработкой ПО. Они требовали, чтобы производители оборудования перестали распространять свои программы бесплатно. В 1969 году суд США признал, что добавление компанией IBM к поставке своего оборудования бесплатного программного обеспечения является нарушением принципов конкуренции. В 1980 году в США авторское право было расширено на компьютерные программы. Свободное развитие ПО было приостановлено.

В 1983 году сотрудник лаборатории искусственного интеллекта Массачусетского технологического института Ричард Столман дал старт проекту GNU, который был нацелен на создание операционной системы и прикладных компьютерных программ, которые не защищены патентами и позволяют пользователям и разработчикам свободно получать, запускать, изучать, изменять и распространять это программное обеспечение.

Сегодня яркими представителями СПО являются веб-браузер Mozilla Firefox, офисный пакет OpenOffice.org, операционные системы на базе ядра Linux. Менее известным среди рядовых пользователей, но критически важным для сети Интернет является такой свободный программный продукт, как веб-сервер Apache, обеспечивающий работу 59% сайтов мира².

Особо актуальным СПО является для академической среды. При проведении математических расчётов учёный всегда может проверить корректность вычисления компьютерной программой того или иного показателя путём анализа исходного кода программы, а не полагаться на «черный ящик», которым является закрытая программа. Пользуясь свободной программой учёный не заставляет своих коллег, которые хотят воспроизвести результаты его работы, тратить деньги на приобретение этой программы, тем самым расширяя круг лиц, которым доступны его выводы. Сохраняя свои тексты, исходные данные и прочие рабочие материалы в свободных электронных форматах, учёный может быть уверен, что его труды будут доступны будущим поколениям. Для закрытых программ и форматов гипотетически возможна такая ситуация, что в будущем будет невозможно открыть тот или иной файл, т.к. фирма, выпускавшая программу обанкротилась и версий этой программы для нового оборудования не создано.

Не меньшие перспективы СПО предоставляет для образования. Информатики могут заглядывать внутрь программы и изучать компьютерную науку на реальных примерах. Каждый студент может получить копию программы, которой его обучают, и это будет совершенно законно. Инициативный студент может включиться в развитие любого, понравившегося ему проекта СПО, развивая свои навыки в предметной области, расширяя свою социальную сеть, улучшая свой английский язык и готовя хороший задел для поиска работы.

СПО для социологических исследований

Представим три свободных программных продукта, используемых непосредственно в процессе полевого социологического исследования.

¹ Free software // Wikipedia http://en.wikipedia.org/wiki/Free_software

² January 2011 Web Server Survey. Netcraft. <http://news.netcraft.com/archives/2011/01/12/january-2011-web-server-survey-4.html>

Интернет-опросы

Всё более популярным становится проведение опросов через Интернет. Существует несколько свободных программ, предназначенных для анкетирования респондентов с помощью веб-технологий. Проект LimeSurvey¹ стартовал в 2003 году в Австралии² и быстро набрал популярность после того, когда в программу были добавлены возможности по вопросам-фильтрам, контролю доступа по приглашениям и созданию шаблонов. Наиболее значимым достижением LimeSurvey стало его использование для сбора сведений о нарушениях на президентских выборах в США в 2004 году. За первые 10 часов голосования с помощью системы LimeSurvey было собрано информации о более чем 13500 нарушениях³. Сейчас центр разработки LimeSurvey находится в Германии.

LimeSurvey русифицирован, он обладает возможностью проводить один и тот же опрос на нескольких языках. Программа позволяет подсчитывать для данных опроса описательную статистику и диаграммы. Для дальнейшего анализа пользователь может экспортировать массив с данными в Excel, SPSS или любую другую программу по статобработке.

Помимо основного предназначения программы мы использовали LimeSurvey для организации ввода данных с бумажных анкет. Работа была распределена между несколькими операторами. Находясь в разных местах, они через Интернет вносили ответы в единую базу данных, что сняло необходимость объединять несколько файлов в один.

Дизайн-макет анкеты

Дизайн анкеты является важным фактором для снижения систематической ошибки наряду с формулировкой текста анкеты и нейтральностью шкал. Существует целая международная ассоциация⁴, в которой аккумулируется опыт по разработке эффективных бланков, в число которых входят и анкеты.

Австралийский социолог Адам Земит, следуя принципам, изложенным в книге⁵ консультанта по бизнес-бланкам Роберта Барнета, члена вышеупомянутой ассоциации, создал свободный формат для подготовки анкет queXML⁶. Использование данного формата позволяет исследователю снять с себя задачу по внешнему оформлению анкеты и полностью сконцентрироваться на её содержании. Особо ценным этот подход становится при использовании анкет на самозаполнение, когда у исследователя нет возможности проводить с респондентом полноценный инструктаж по работе с анкетой.

Используя формат queXML, исследователь автоматически оформляет свою анкету в соответствии с современными мировыми наработками в области дизайна бланков, тем самым повышая качество получаемых ответов. В queXML учитывается тип и размер шрифта, толщина обрамляющих линий, использование «путеводных» линий для взгляда респондента и пр. Мы использовали формат queXML в создании анкет для опросов в секторе B2B.

Формат queXML является частью свободного пакета queX Suite⁷, который также включает в себя систему по сканированию и распознаванию бумажных анкет, систему для проведения телефонных опросов и систему для очистки и кодирования анкетных данных.

Статистический анализ данных

В апреле 2008 года компания SPSS выпустила шестнадцатую версию своего статистического пакета. Среди ряда обновлений в программу была добавлена поддержка языка R, что позволило пользователям SPSS «получить доступ к множеству статистических алгоритмов,

¹ LimeSurvey - the free & open source survey software tool <http://www.limesurvey.org>

² LimeSurvey // Wikipedia <http://en.wikipedia.org/wiki/LimeSurvey#History>

³ PHSurveyor Plays Key Role in US Elections. Novemer 6th, 2004 VerifiedVoting.org, Inc. <http://www.verifiedvoting.org/article.php?id=5207>

⁴ The Association for Forms and Business Process Professionals <http://www.bfma.org/>

⁵ Barnett R. The Form Designer's Quick Reference Guide. Robert Barnett & Associates Pty, 1994.

⁶ queXML An Open Source, XML, multi-mode questionnaire description toolkit <http://quexml.sourceforge.net/>

⁷ queX Suite <http://quexml.sourceforge.net/node/14>

созданных в R, и использовать их в SPSS как часть синтаксиса¹». Что заставило разработчиков такого мощного пакета как SPSS включить поддержку некоего стороннего продукта?

R — это язык программирования, а также программная среда для статистического анализа и работы с графикой. R используется «растущим числом аналитиков в бизнесе и в науке. Он становится их общепринятым языком²». Среди корпоративных пользователей R такие компании как Google, Facebook, Pfizer, Bank of America.

Начало проекта было положено Росом Айхэкой и Робертом Джентельменом в университете Окленда (Новая Зеландия)³. Сейчас R предоставляет пользователю массу методов, включая линейное и нелинейное программирование, анализ временных рядов, классификацию и кластеризацию и пр. Преимущество R состоит в том, что это не монолитная программа, а ядро, расширяемое за счёт дополнений. Создание новых дополнений не требует знания никаких языков программирования кроме самого R. Это позволяет пользователям оперативно добавлять в среду R новейшие статистические методики и сразу же распространять это дополнение в сообществе. На конец января 2011 года в библиотеке R было зарегистрировано 2787 дополнений⁴.

Существует целый ряд проектов по созданию графического интерфейса для R, чтобы приблизить работу в R к принципам хорошо знакомых пакетов наподобие SPSS. Несмотря на это, вся мощь R будет доступна только тем исследователям, которые изучат этот язык программирования и смогут работать через синтаксис. Дополнительную сложность в России составляет незначительное число литературы про R на русском языке.

Заключение

Свободное программное обеспечение несёт в себе широкие возможности для науки и образования. На сегодняшний день практически для любой задачи, с которой сталкивается исследователь, можно найти прикладную программу, распространяемую по условиям свободных лицензий. Пользователь может получить эту программу, использовать её для любых, в том числе коммерческих, целей, ознакомиться с её принципами действия, изменить эту программу под свои специфические требования, и распространить оригинальную или модифицированную версию в любых количествах.

О влиянии опыта интервьюера на результаты телефонного интервью

Проскурина Дарья Андреевна
*Ивановский государственный энергетический
университет им. В.И. Ленина*

В литературе, посвященной проблеме эффекта интервьюера, в качестве одного из факторов, вызывающих ответные смещения, выступает опыт интервьюера. В целом изучению искажающего воздействия этого фактора посвящено очень немного работ, при этом полученные результаты неоднозначны⁵. В частности, Т. ван Тилбург обнаружил, что опыт, накапливаемый интервьюерами в течение одного исследования, действительно оказывает влияние на результаты их работы, в то время как Т. Джонсону с коллегами не удалось зафиксировать взаимосвязь между опытом полевых работников и самоотчетами респондентов об употреблении наркотиков. Э. Сингер и ее коллеги пришли к выводу о том, что наличие некоторого опыта у интервьюера лучше, нежели его отсутствие, однако после определенного момента

¹ What's new in SPSS 16.0. 2007 SPSS Inc.

² Vance A. Data Analysts Captivated by R's Power // New York Times. 2009. January 6.

³ R (programming language) [http://en.wikipedia.org/wiki/R_\(programming_language\)](http://en.wikipedia.org/wiki/R_(programming_language))

⁴ Contributed Packages. <http://cran.r-project.org/web/packages/>

⁵ См., например: Johnson T., Fendrich M., Shaligram Ch. et al. An evaluation of the effects of interviewer characteristics in an RDD telephone survey of drug use // Journal of Drug Issues. 2000. Vol. 30. № 1. P. 77–101; Singer E., Frankel M., Glassman M. The effect of interviewer characteristics and expectations on response // Public Opinion Quarterly. 1983. Vol. 47. № 1. P. 68–83; Tilburg van T. Interviewer effects in the measurement of personal network size: A Nonexperimental study // Sociological Methods and Research. 1998. Vol. 26. № 3. P. 300–328.

увеличение опыта ведет к снижению уровня отвечаемости¹. А.Ю. Мягков и И.В. Журавлева обнаружили, что опытные интервьюеры тратили на проведение персональных интервью в среднем чуть больше времени, чем они ожидали до начала полевой работы, в то время как новички оказались точнее в своих прогнозах².

Когда опыт интервьюера определяется как его общий опыт участия в опросах, у опытных интервьюеров, как правило, наблюдается более высокое качество данных (ниже процент пропусков ответов и социально желательных ответов на сенситивные вопросы, более значимые корреляции между ключевыми переменными исследования и т.д.)³. Ч. Кеннелл и его коллеги утверждают, что темп беседы влияет на внимание, уделяемое респондентом вопросу. Более быстрый темп интервью может привести к снижению качества ответов хотя бы потому, что респондент не может уделить достаточно времени всем этапам процесса формирования ответа⁴.

Неоднозначна трактовка самого понятия «опыт интервьюера». На сегодняшний день можно выделить несколько различных подходов к его определению и фиксации:

- опыт работы интервьюером, фиксируемый посредством общего количества исследований, в которых полевой работник когда-либо принимал участие⁵;
- общий опыт как показатель временной продолжительности работы интервьюером, измеряемый в месяцах, годах и т.п.⁶;
- опыт работы в одном опросном цикле, накапливаемый по мере проведения интервью и фиксируемый с помощью порядкового номера проводимой беседы⁷;
- опыт как синоним квалификации интервьюера, показателем которой выступает качество полевой работы интервьюера, оцениваемое исследователем по ряду критериев после того, как интервью завершено⁸;
- опыт участия в разных этапах одного лонгитюдного исследования⁹;
- опыт работы в конкретной опросной организации¹⁰.

Проверка гипотезы о влиянии опыта интервьюеров на ответы респондентов и качество полевой работы являлась одной из целей нашего методического эксперимента, проведенного нами в 2010г. в режиме телефонного интервью на базе одной из опросных организаций г. Иванова под руководством доктора социологических наук, профессора А.Ю. Мягкова. В эксперименте приняли участие 20 интервьюеров, каждый из которых провел в общей сложности 40 стандартизированных интервью по специально разработанному вопроснику, включавшему среди прочих блок сенситивных вопросов (5 вопросов о воровстве, употреблении

¹ *Singer E., Frankel M., Glassman M.* Op. cit. P. 80.

² См.: *Мягков А.Ю., Журавлева И.В.* Эффект ожиданий интервьюера в персональном интервью // Социологический журнал. 2004. № 3/4. С. 14.

³ См., например: *Singer E., Frankel M., Glassman M.* Op.cit. P. 68–83; *O’Muircheartaigh, C., Campanelli, P.* The relative impact of interviewer effects and sample design effects on survey precision // *Journal of the Royal Statistical Society. Series A.* Vol. 161. № 1. P. 63–77.

⁴ *Cannell C.F., Miller P.V., Oksenberg L.* Research on Interviewing Techniques // *Sociological Methodology*, 1981 / Ed. by S. Leinhardt. San Francisco: Jossey-Bass, 1981. P. 107–125.

⁵ *Olson K., Peytchev A.* Effect of interviewer experience on interview pace and interviewer attitudes // *Public Opinion Quarterly*. 2007. Vol. 71. № 2. P. 273–286.

⁶ См., например: *Groves R.M., Fultz N.H.* Gender effects among telephone interviewers in a survey of economic attitudes // *Sociological Methods and Research*. 1985. Vol. 14. № 1. P. 36; *Johnson T., Fendrich M., Shaligram Ch.* et al. Op. cit. P. 83; *Kane E.W., Macaulay L.J.* Interviewer gender and gender attitudes // *Public Opinion Quarterly*. 1993. Vol. 57. № 1. P. 6; *Singer E., Frankel M., Glassman M.* Op. cit. P. 69.

⁷ См., например: *Olson K., Peytchev A.* Op. cit. P. 274.

⁸ См., например: *Freeman J., Butler E.* Some sources of interviewer variance in surveys // *Public Opinion Quarterly*. 1976. Vol. 40. № 1. P. 79–91.

⁹ См., например: *Lipps O.* Interviewer and respondent survey quality effects in a CATI panel // *Bulletin de Méthodologie Sociologique*. 2007. Vol. 95. № 1. P. 5–25; *Lord V.B., Friday P.C., Brennan P.K.* The effects of interviewer characteristics on arrestees’ responses to drug-related questions // *Applied Psychology in Criminal Justice*. 2005. Vol. 1. № 1. P. 36–55.

¹⁰ См., например: *Olson K., Peytchev A.* Op. cit. P. 274.

наркотиков и алкоголя, супружеской измене и сексуальных отношениях). Общая выборка в эксперименте, таким образом, составила 800 респондентов.

Опыт интервьюера фиксировался двумя способами: как общее количество исследований, в которых интервьюер когда-либо принимал участие, и как опыт работы сотрудника в конкретной опросной организации, оцениваемый экспертами. Таким образом, в соответствии с первым критерием интервьюеры были разделены на 4 категории: (1) интервьюеры, не имеющие опыта проведения интервью (0 исследований); (2) интервьюеры, имеющие незначительный опыт (1–2 исследования); (3) интервьюеры, имеющие средний опыт (3–10 исследований); (4) интервьюеры, имеющие большой опыт (более 10 исследований). В соответствии со вторым критерием были выделены три группы интервьюеров: (а) новички (интервьюеры, работающие в данной опросной организации впервые), (б) эпизодические участники (интервьюеры, участвующие в опросах данной фирмы, от случая к случаю), (в) постоянные интервьюеры (полевые работники, принимающие участие во многих исследованиях, проводимых данной опросной организацией).

В ходе эксперимента мы изучали влияние опыта интервьюеров на длительность интервью, число отказов от участия в исследовании и количество искренних ответов на вопросы сенситивного блока.

Опыт интервьюера и длительность интервью

Длительность интервью в нашем эксперименте оценивалась как по объективным, так и по субъективным показателям. Объективная длительность каждого интервью рассчитывалась на основании фиксации интервьюером времени начала и завершения каждого интервью. Субъективная длительность оценивалась на основании ответов респондентов на завершающий каждую беседу вопрос о том, сколько минут, по их ощущениям, длилось данное интервью. Помимо этого до начала полевого этапа мы измеряли ожидания интервьюеров относительно длительности предстоящих интервью.

Результаты исследования показывают, что более опытные интервьюеры проводили интервью быстрее, чем их менее опытные коллеги (табл. 1), однако различия статистически не значимы. Интервьюеры, оцениваемые опросной организацией как постоянные участники исследований, тратили на одну беседу в среднем 8,16 минут, в то время как «эпизодическим» участникам опросов требовалось 8,62 минуты, а «новичкам» 8,69 минуты.

Таблица 1

Влияние опыта интервьюера, фиксируемого экспертным способом, на показатели длительности интервью

Опыт интервьюера	Объективная длительность, мин.	Субъективная длительность, мин.	Разность между объективной и субъективной длительностью, мин.	Ожидания длительности, мин.
«Новички»	8,69	7,55	1,14	9,00
«Эпизодические участники»	8,62	7,75	0,87	11,67
Постоянные участники	8,16	7,37	0,79	12,42

Показательно, что предварительные ожидания длительности предстоящих интервью среди новичков оказывались в среднем значительно более оптимистичными, нежели среди их более опытных коллег. Так, «новички» перед началом полевой работы полагали, что в среднем проведение одного интервью будет занимать у них 9,00 минут, в то время как «эпизодические» участники рассчитывали справиться с проведением интервью за 11,67 минут ($t=8,753$, $p=0,000$), а постоянные интервьюеры за 12,42 минут ($t=9,652$, $p\leq 0,000$).

В качестве одного из показателей коммуникативной успешности интервьюера можно считать разность между объективной и субъективной длительностью интервью: чем меньше субъективная длительность интервью, тем более успешно интервьюер справляется с комму-

никативной задачей. Однако зафиксированные различия по этому показателю также оказались статистически не значимы.

Аналогичные тенденции мы наблюдаем и в том случае, когда опыт полевой работы фиксировался через общее количество исследований, в которых сотрудник принимал участие (табл. 2). Так, опытные интервьюеры (11 и более исследований) тратили на проведение одного интервью в среднем 7,91 минут, в то время как полевым работникам со средним опытом интервьюирования (3–10 исследований) требовалось в среднем 8,61 минут, а интервьюерам с небольшим опытом (1–2 исследования) — 9,52 минут. Следует, однако, отметить, что полевые работники, совсем не имеющие опыта интервьюирования, тратили на проведение одной беседы в среднем 8,30 минут, что несколько меньше, чем среди их коллег с небольшим и средним опытом. Выявленные различия и тут статистически не значимы.

Таблица 2

Влияние опыта интервьюера, фиксируемого количественным способом, на показатели длительности интервью

Опыт интервьюера	Объективная длительность, мин.	Субъективная длительность, мин.	Разность между объективной и субъективной длительностью, мин.	Ожидания длительности, мин.
Нет опыта	8,30	7,62	0,68	9,17
Небольшой опыт (1–2 исследования)	9,52	8,12	1,40	11,67
Средний опыт (3–10 исследований)	8,61	7,67	0,94	10,69
Большой опыт (11 и более исслед.)	7,91	6,99	0,92	11,90

Что касается разности между объективной и субъективной длительностью интервью, наиболее успешными нам представляются интервьюеры, имеющие опыт проведения 1–2 опросов (показатель в этой группе интервьюеров равен 1,40 минуты), в то время как наименьшее значение разности получено интервьюерами без полевого опыта (0,68 минуты).

Полученные результаты отчасти совпадают с выводами, описанными в специальной литературе. Ряд исследователей отмечают, что опытные интервьюеры тратят значительно меньше времени на проведение интервью, нежели новички¹. Э. Блэйр объясняет подобные эффекты тем, что интервьюеры, давно участвующие в опросах, менее строго следуют предписанным правилам интервьюирования². Они чаще допускают отступления от вопросных формулировок, не полностью их прочитывают, пропускают отдельные слова, фразы и даже целые вопросы, ответы на которые кажутся им очевидными, исходя из общего контекста интервью и собственных ситуативных ожиданий³.

Опыт интервьюера и число отказов от участия

Что касается коммуникативной успешности, то, как свидетельствуют результаты нашего эксперимента, опытные полевые работники проявили себя значительно более эффективными по сравнению с остальными своими коллегами (табл. 3). Интервьюеры, квалифицированные опросной организацией как постоянные, получили в среднем 18,5 отказов от участия в исследовании со стороны респондентов, в то время как эпизодические участники – 21,09 отказа ($t=2,295$, $p \leq 0,05$), а новички — 22,38 отказов ($t=3,349$, $p \leq 0,001$).

При фиксации опыта посредством общего количества исследований, в которых интервьюер когда-либо принимал участие, картина выглядит несколько иначе (табл. 4). Наименьшее число отказов было получено интервьюерами, никогда не участвовавшими в опросах (в среднем 16,25); с наибольшим количеством коммуникативных неудач столкнулись полевые

¹ Groves R.M., Fultz N.H. Gender effects among telephone interviewers in a survey of economic attitudes // Sociological Methods and Research. 1985. Vol. 14. № 1. P. 39; Мягков А.Ю., Журавлева И.В. Указ. соч. С. 78–107.

² Bradburn N.M., Sudman S. Improving Interview Method and Questionnaire Design. San Francisco: Jossey-Bass, 1979. P. 49.

³ Ibid. P. 29.

работники, имеющие опыт проведения одного-двух опросов (в среднем 29,00, $t=7,390$, $p=0,000$). При этом интервьюеры со средним и большим опытом полевой работы сталкивались в ходе исследования с 19,33 и 21,80 отказами, соответственно (выявленные различия статистически не значимы).

Таблица 3

Влияние опыта интервьюера, фиксируемого экспертным способом, на число отказов от участия

Опыт интервьюера	Число отказов, <i>абс. числа</i>
Новички	22,38
Эпизодические участники	21,09
Постоянные участники	18,5

Таблица 4

Влияние опыта интервьюера, фиксируемого количественным способом, на число отказов от участия

Опыт интервьюера	Число отказов, <i>абс. числа</i>
Нет опыта	16,25
Небольшой опыт (1–2 исследования)	29,00
Средний опыт (3–10 исследований)	19,33
Большой опыт (11 и более исследований)	21,80

Опыт интервьюера и количество искренних ответов на чувствительные вопросы

Одной из ключевых составляющих нашего опросного инструментария был блок чувствительных вопросов, в который вошли 5 поведенческих вопросов о воровстве, употреблении наркотиков и алкоголя, супружеской измене и сексуальных отношениях. При выборе чувствительной проблематики в качестве базовой для выявления интересующих нас эффектов мы руководствовались мнением о том, что общение на деликатные темы сделает эффект интервьюера более выраженным и легко наблюдаемым¹. Анализируя ответы респондентов на чувствительные вопросы, мы исходили из предположения, что искренними признаниями можно считать утвердительные ответы на вопросы о социально неодобряемом или стигматизированном поведении. В ходе анализа влияния опыта интервьюера, измеренного экспертным способом, мы обнаружили, что интервьюеры, квалифицированные организаторами исследования как «новички», получили наименьшее число признаний на чувствительные вопросы по сравнению с их более опытными коллегами (в среднем 1,13 признаний из пяти возможных, табл. 5). При этом наивысший показатель – 1,33 признания из пяти возможных – зафиксирован среди интервьюеров, принимающих участие в опросах организации от случая к случаю ($t=2,091$, $p\leq 0,05$).

Таблица 5

Влияние опыта интервьюера, измеряемого экспертным способом, на количество признаний респондентов на чувствительные вопросы

Опыт интервьюера	Количество признаний респондентов на чувствительные вопросы, <i>абс. числа</i>
Новички	1,13
Эпизодические участники	1,33
Постоянные участники	1,27

В случае, когда опыт полевых работников фиксировался как общее число исследований, в которых они принимал участие, максимальное число признаний по блоку чувствительных вопросов получили интервьюеры с небольшим опытом (1–2 исследования): они получали в среднем 1,60 искренних признаний по чувствительным вопросам из 5 возможных, (табл. 6). Наименьшее число признаний — 1,07 из предельно возможных 5 — в среднем получили ин-

¹ См., например: Мязков А.Ю., Журавлева И.В. Указ. соч. С. 83.

тервьюеры с опросным опытом 3–10 исследований ($t=2,387$, $p\leq 0,05$). Какой-либо четкой статистически значимой взаимосвязи между переменной опыта, фиксируемой количественно, и числом признаний по чувствительным вопросам не выявлено.

Таблица 6

Влияние опыта интервьюера, измеряемого количественным способом, на число признаний респондентов на чувствительные вопросы

Опыт интервьюера	Количество признаний респондентов на чувствительные вопросы, <i>абс. числа</i>
Нет опыта	1,19
Небольшой опыт (1–2 исследования)	1,60
Средний опыт (3–10 исследований)	1,07
Большой опыт (11 и более исследований)	1,31

В ходе проведенного нами эксперимента мы пришли к выводу, что опыт интервьюера действительно оказывает влияние на результаты работы и качество опросных данных в телефонном интервью. Так «новички» в среднем тратили на проведение одного интервью больше времени, чем их опытные коллеги, а также получали по сравнению с ними значительно большее число отказов. При этом наименьшая разность между объективной и субъективной длительностью интервью у «новичков» может свидетельствовать в пользу их большей коммуникативной успешности во время интервью по сравнению с более опытными коллегами.

Безусловно, влияние опыта интервьюера на качество опросных данных не исчерпывается лишь рассмотренными нами аспектами работы интервьюера (длительность интервью, уровень отказов и количество искренних признаний на чувствительные вопросы). Не исключено, что фиксация опыта интервьюера иными способами приведет и к иным результатам, которые могут лечь в основу рекомендаций для практикующих социологов, и впоследствии использоваться на практике в зависимости от поставленных исследовательских целей и задач. Необходимо дальнейшее детальное изучение данной проблематики с использованием методических экспериментов.