

МЕТОДЫ И МЕТОДОЛОГИЯ

DOI: 10.14515/monitoring.2017.5.07

Правильная ссылка на статью:

Мавлетова А. М. Использование параданных в опросах для корректировки и оптимизации полевых работ в адаптивном дизайне // Мониторинг общественного мнения: Экономические и социальные перемены. 2017. № 5. С. 105—119. DOI: 10.14515/monitoring.2017.5.07.

For citation:

Mavletova A. M. Using survey paradata for the fieldwork corrections and optimization in adaptive design. *Monitoring of Public Opinion: Economic and Social Changes*. 2017. № 5. P. 105—119. DOI: 10.14515/monitoring.2017.5.07.

А. М. Мавлетова

ИСПОЛЬЗОВАНИЕ ПАРАДАНЫХ В ОПРОСАХ ДЛЯ КОРРЕКТИРОВКИ И ОПТИМИЗАЦИИ ПОЛЕВЫХ РАБОТ В АДАПТИВНОМ ДИЗАЙНЕ

ИСПОЛЬЗОВАНИЕ ПАРАДАНЫХ В ОПРОСАХ ДЛЯ КОРРЕКТИРОВКИ И ОПТИМИЗАЦИИ ПОЛЕВЫХ РАБОТ В АДАПТИВНОМ ДИЗАЙНЕ

USING SURVEY PARADATA FOR THE FIELDWORK CORRECTIONS AND OPTIMIZATION IN ADAPTIVE DESIGN

МАВЛЕТОВА Айгуль Маратовна — кандидат социологических наук, доцент, Департамент социологии; старший научный сотрудник, Международная лаборатория сравнительных социальных исследований, Национальный исследовательский университет «Высшая школа экономики», Москва, Россия.

E-MAIL: amavletova@hse.ru

ORCID: 0000-0001-7407-2072

Aigul M. MAVLETOVA¹ — Cand. Sci (Sociol.), Associate Professor, Senior Research Fellow

E-MAIL: amavletova@hse.ru

ORCID: 0000-0001-7407-2072

¹ National Research University Higher School of Economics, Moscow, Russia

Аннотация. Параданные — данные о процессе сбора данных (наблюдения интервьюеров; данные о количестве визитов или звонков в домохозяйство; время заполнения анкеты и т. д.). В статье рассматриваются возможности использования параданных в рамках адаптивного дизайна, отдельные элементы которого по ходу выполнения полевых работ могут варьироваться

Abstract. Paradata are data about the data collection process (interviewers' observations, number of phone calls or visits to a household, time spent completing the questionnaire, etc.). The article considers the opportunities to use paradata in adaptive design where separate elements may vary depending on various categories of respondents during the fieldwork. Researchers

для разных категорий респондентов. Исследователем при этом учитывается и стоимость исследования, и ошибка неответов. Анализ параданных показал, что оценки и наблюдения интервьюеров (описание состава района и дома, оценка вероятности, что домохозяйство примет участие в опросе) могут быть достаточно точными, что позволяет использовать их в адаптивном дизайне. Данные о дате, времени и количестве контактов позволяют планировать оптимальный день и время визита или телефонного звонка интервьюера; устанавливать приоритет некоторым категориям респондентов для увеличения их отклика в ходе полевых работ. Рассмотрены недостатки и сложности использования параданных в адаптивном дизайне.

Ключевые слова: параданные, адаптивный дизайн исследования, телефонные опросы, личные опросы, интервьюер, наблюдения интервьюеров, бланк контактов

Благодарность. Статья подготовлена при поддержке Российского научного фонда (Грант № 17-78-20172).

should take into consideration both the fieldwork costs and nonresponse bias. The paradata analysis shows that interviewers' assessments and observations (description of the district and house, assessment of the likelihood of the household's participation in the survey) can be accurate enough to be used in adaptive design. Date, time and number of contacts help to predict the optimal day and time for the interviewers' contact; researcher can prioritize certain categories of respondents to increase their response rate during the fieldwork. The author also considers shortcomings and difficulties in using the paradata in adaptive design.

Keywords: paradata, adaptive survey design, CATI, face-to-face surveys, interviewer, interviewers' observations, contact form

Acknowledgment. The paper is supported by the Russian Science Foundation under grant no. 17-78-20172.

Введение

С началом активного применения компьютерных технологий в области социологических исследований наряду с термином «данные» были введены термины «параданные» [Cooper, 1998], и метаданные [Cooper, 2005]. Метаданные — это информация о данных, которая включает описание выборки и использованных методов сбора данных, структуры собранных данных (переменные и значения переменных в файлах данных), инструментария, а также информацию о получении доступа к данным [Сидоров, 2011]. Если метаданные — это данные о данных,

то параданные — это данные о разных процессах их сбора: о времени заполнения опросных листов, о количестве и времени визитов или звонков в домохозяйство и др. [Couper, 2005].

Несмотря на огромное количество англоязычных публикаций об использовании параданных в разных типах опросов, нет стандартного определения этого понятия или стандартной классификации параданных. Кроме того, не всегда можно провести четкое различие между метаданными и параданными. Ф. Кройтер для простоты разделяет метаданные и параданные по принципу данных макро- и микроуровня [Kreuter, 2013]. Соответственно, метаданные — это данные макроуровня, которые позволяют понять общую структуру данных, в то время как параданные — это данные микроуровня, которые создаются непосредственно в процессе сбора данных. Можно привести и связь между метаданными и параданными: так, например, на основе среднего времени заполнения на уровне респондента/интервьюера (параданные) можно рассчитать среднее время заполнения анкеты в целом (метаданные) [Kreuter, 2013]. Между тем, если в области метаданных распространяются стандарты их сбора и представления¹, то в области параданных подобной инициативы пока нет.

Интерес к параданным растет не только у западных, но и у отечественных исследователей [Ипатова, 2014, 2016; Рогозин, Ипатова, 2012, 2015; Сидоров, 2011]. Мы нашли две обзорные статьи по использованию параданных в отечественных журналах [Ипатова, 2014; Сидоров, 2011]. Н. Сидоров подчеркивает как специфику параданных в зависимости от методов исследования, так и особенности их использования в зависимости от цели исследователя [Сидоров, 2011]. А. Ипатова рассматривает применение параданных в телефонных опросах и показывает возможности их использования на примере информации о звонках: о дате и времени начала и завершения интервью, длительности интервью, длительности звонка, а также о результате соединения и результате дозвона [Ипатова, 2014].

В нашем обзоре мы сфокусируем внимание на том, как можно использовать параданные для внесения изменений в дизайн исследования по ходу полевых работ в личных и телефонных опросах. Применение параданных для оптимизации полевых работ является частью так называемого адаптивного дизайна исследования (adaptive или responsive survey design).

Структура статьи следующая. Сначала мы рассмотрим элементы адаптивного дизайна, возможности применения параданных для оптимизации полевых работ в рамках адаптивного дизайна, после чего обсудим возможные недостатки и сложности использования параданных в адаптивном дизайне. В заключение сделаем выводы.

Адаптивный дизайн исследования

В последние годы в методологии социологических исследований растет внимание к обсуждению адаптивного дизайна [Axinn, Link, Groves, 2011; Groves, Heeringa, 2006; Schouten, Calinescu, Luiten, 2013; Schouten, Peytchev, Wagner, 2017; Tourangeau et al., 2017; Wagner, 2008]. Р. Грувз и С. Херинга вводят понятие

¹ См., например, <http://www.ddialliance.org/>.

адаптивного дизайна исследования, который приходит на смену традиционному дизайну [Groves, Heeringa, 2006]. Если в традиционном дизайне исследования основное внимание уделялось уменьшению ошибок опроса, то в адаптивном дизайне исследователь учитывает как ошибки опроса, так и стоимость исследования, пытаясь найти оптимальное решение, исходя из имеющегося бюджета. Безусловно, переход к адаптивному дизайну вызван, с одной стороны, уменьшением доли откликов в опросах, с другой стороны, сокращением бюджетов исследований. Как правило, адаптивный дизайн направлен либо на общее увеличение доли откликов, либо на выравнивание доли откликов между разными категориями, либо на баланс выборки по определенным социально-демографическим переменным, но с учетом стоимости исследования.

Адаптивный дизайн может иметь разные формы. Например, он может быть направлен на разработку разных подходов к респондентам в зависимости от их характеристик, на их стратификацию в зависимости от вероятности участия в опросе и установление приоритета категориям респондентов с менее или, наоборот, более высокой вероятностью участия (case prioritization), а также на разработку стратегии по привлечению к участию в исследовании определенных труднодоступных категорий респондентов. Кроме того, в рамках адаптивного дизайна можно вычислять оптимальное время для завершения полевых работ или для изменения или ввода дополнительного метода опроса [Tourangeau et al., 2017].

Р. Гривз и С. Херинга описывают следующие шаги адаптивного дизайна [Groves, Heeringa, 2006]:

1. предварительно определите элементы дизайна исследования, потенциально влияющие как на ошибки, так и на стоимость;
2. определите основные индикаторы стоимости и ошибок исследования;
3. следите за этими индикаторами на начальных этапах процесса сбора данных;
4. скорректируйте элементы дизайна исследования на последующих этапах сбора данных, находя компромиссное решение между ошибками опроса и стоимостью исследования;
5. объедините данные разных этапов процесса сбора данных в единую оценку.

Изначально исследователи выбирают определенные элементы дизайна, которые одновременно влияют на ошибки опроса и его стоимость (например, метод опроса, метод рекрутинга, дизайн выборки, количество визитов и т. д.). В рамках адаптивного дизайна можно выделить два основных направления. Представим это на примере. Предположим, что исследователь выбрал вознаграждение и количество визитов в домохозяйство как основные элементы адаптивного дизайна. При этом наибольшее внимание исследователь уделяет различию в доле откликов между разными категориями респондентов. В рамках первого направления разные домохозяйства получили бы разное вознаграждение и разное количество визитов интервьюера, то есть элементы дизайна исследования варьировались в зависимости от характеристик респондентов (например, такую интерпретацию адаптивного дизайна см. [Schouten, Calinescu, Luiten, 2013]). Таким образом, исследователь бы выбрал стратегию увеличения количества визитов и размера вознаграждения в определенные домохозяйства, в которых предполагается низкая вероятность участия (например, в больших городах исследователь может

выбрать и больший размер вознаграждения, и большее количество визитов). В рамках второго направления элементы дизайна меняются по ходу выполнения полевых работ. Например, по мере проведения исследования может быть принято решение и об увеличении вознаграждения всем или определенным категориям участников, и об изменении максимального количества визитов в определенные домохозяйства. Подобные изменения могут происходить несколько раз в ходе полевых работ. Поскольку дизайн исследования меняется по ходу полевых работ, иногда в литературе его называют динамичным [Tourangeau et al., 2017]. Наибольшее внимание исследователей в настоящее время приковано ко второму направлению в рамках адаптивного дизайна; именно о такой интерпретации адаптивного дизайна писали Р. Грувз и С. Херинга [Groves, Heeringa, 2006].

Для оптимизации дизайна исследования используются разные данные, включая базовые данные о первичной единице отбора (такие как социально-демографический профиль, численность населения и другие возможные данные), данные из прошлых волн исследования в том случае, если это повторяющееся или лонгитюдное исследование, а также параданные. По факту, если данные о первичной единице отбора и данные из прошлых волн у исследователя есть на руках до начала проведения полевых работ для построения дизайна, то параданные появляются лишь в ходе проведения опроса. Особое место в адаптивном дизайне занимает именно использование параданных, так как на основе этих данных о полевых работах исследователь может принимать решение об изменениях в дизайне опроса. Предполагается, что исследователь должен получать и анализировать информацию на ежедневной основе, чтобы принимать решение об изменениях в последующих этапах сбора данных. При этом необходимо следить как за общей долей откликов и откликов в определенных категориях, так и за переменными, которые были выбраны в качестве ключевых для учета баланса выборки [Kirgis, Lepkowski, 2013].

В адаптивном дизайне прежде всего используются два типа параданных, на основе которых можно следить за полевыми работами, анализировать имеющиеся данные и принимать решение об изменениях в дизайне опроса:

- наблюдения и оценки интервьюеров;
- данные о контактах (дата, время и количество контактов).

Далее кратко опишем возможности их использования в адаптивном дизайне.

Наблюдения и оценки интервьюеров

В случае проведения опроса методом личного интервью у респондента дома интервьюеров могут просить заполнить карточку наблюдений: например, описать дом, двор или округ, а также оценить, кто может проживать по адресу. Такие данные могут использоваться для косвенной оценки ошибки неответов, оценки вероятности участия домохозяйства в опросе, а также оценки усилий, которые необходимы исследователю для получения ответа от того или иного домохозяйства. Подобные данные можно получить как до начала проведения основного интервью, так и в ходе проведения полевых работ.

В известном американском исследовании «*National Survey of Family Growth*», которое активно использует параданные для оптимизации полевых работ, интер-

вьюер должен описать район (жилые/нежилые строения) и дом (многоэтажный многоквартирный жилой дом, дом на одну семью и т. д.), предположить наличие жителей, не говорящих по-английски, оценить безопасность района и социально-демографический профиль домохозяйства с фокусом на наличие детей (наличие колясок, детских игрушек, детских велосипедов и т. д.) [Wagner, 2008]. В британском панельном исследовании «Understanding Society» интервьюеров просили указать тип жилища, количество этажей в доме, в случае многоэтажного дома — этаж, на котором находится квартира, наличие преград для попадания в дом, наличие заброшенного сада/огорода возле дома. Интервьюеры должны были оценить вероятность наличия в домохозяйстве детей в возрасте до десяти лет и машины, описать район: наличие заброшенных зданий и мусора на улицах, дать среднюю оценку состояния домов, состояния выбранного дома по сравнению с другими домами этого района [Sturgis, Brunton-Smith, 2011]. Кроме того, подобные наблюдения интервьюеров используются в больших международных исследованиях, например, в Европейском социальном исследовании [Blom, 2009].

Многие исследователи указывают на то, что оценки и наблюдения интервьюеров могут быть достаточно точными, что позволяет использовать их в адаптивном дизайне. Так, Б. Уэст обнаружил, что в 72 % случаях интервьюеры правильно указывали на возможное наличие или отсутствие детей в домохозяйстве [West, 2013]. При этом процент точности оказался ниже в случае наличия ребенка в домохозяйстве, то есть интервьюеру труднее предсказать наличие ребенка дома, чем его отсутствие. Кроме того, интервьюеры могут с достаточной точностью предсказать вероятность участия домохозяйства в исследовании. Р. Гривз и С. Херинга отмечают, что в тех домохозяйствах, вероятность участия которых интервьюеры оценивали как низкую, доля откликов в итоге оказалась ниже (доля откликов 38,5 %) по сравнению с теми домохозяйствами, вероятность участия которых интервьюеры оценивали как более высокую (доля откликов 73,7 %) [Groves, Heeringa, 2006].

В исследовании, посвященном оценке безработицы в Германии, в 73 % случаях интервьюеры правильно предсказывали наличие в домохозяйстве получающих пособия по безработице. В 78 % случаях интервьюеры правильно предсказывали отсутствие в домохозяйстве получающих пособия по безработице [Sinibaldi, Trappmann, Kreuter, 2014]. Кроме того, интервьюеры достаточно точно угадывали низкий доход домохозяйства (82 %), в то время как точность оценки среднего и высокого дохода оказалась ниже. Сравнение данных наблюдений интервьюеров с коммерческими данными о доме и социально-демографическом профиле домохозяйства (наличие молодых людей до 30 лет, тип семьи, оценка дохода домохозяйства) показало более высокую точность предсказаний интервьюеров [Sinibaldi, Trappmann, Kreuter, 2014].

В том случае, если контакт с кем-то из членов домохозяйства состоялся, можно получить более точные оценки интервьюера о том, какое домохозяйство может проживать по этому адресу. При каждом контакте интервьюер может указывать дополнительную информацию: например, были ли заданы домохозяевами какие-то вопросы; было ли сказано, что сейчас неподходящее время для опроса; были ли даны негативные оценки опросу и т. п. [Groves, Heeringa, 2006]. Подобные наблю-

дения интервьюеров могут быть использованы как в личных, так и в телефонных опросах. Когда домохозяйство отказывается принять участие в исследовании, интервьюер может оценить причину отказа. Кроме того, наблюдения интервьюеров используются и для дополнительной оценки ошибки измерения, когда интервьюеры оценивают, насколько, по их мнению, были понятны для респондентов вопросы, насколько они были уверены в своих ответах и т. д.

Каким образом данные наблюдения можно использовать для корректировки и оптимизации полевых работ в рамках адаптивного дизайна? Р. Грувз и С. Херинга показали, что такие характеристики, как несогласие участвовать в опросе в предыдущих визитах или высказывания возражений против участия в опросе, а также затруднения попасть в дом или квартиру значительно снижали вероятность проведения интервью в последующих визитах [Groves, Heeringa, 2006]. Исследователи проанализировали параданные для отбора домохозяйств, не принявших участие в опросе на первом этапе. Для сокращения расходов на полевые работы были отобраны все адреса с высокой вероятностью участия в исследовании и 50 % адресов с более низкой вероятностью участия в исследовании [Groves, Heeringa, 2006]. Данная стратегия наряду с другими дополнительными элементами дизайна позволила увеличить общую долю отклика и частично снизить затраты на проведение полевых работ.

Поскольку оценки интервьюеров показали свою эффективность в личных интервью, подобные данные стали применять и в телефонных опросах [Eckman, Sinibaldi, Möntmann-Hertz, 2013; Sinibaldi, Eckman, 2015]. Так, включение оценки интервьюером вероятности того, что респондент ответит на анкету при следующем телефонном звонке, увеличивает точность статистической модели по сравнению с моделями, в которых экстраполируются только данные о предыдущих контактах: количество звонков, дата, время и диспозиционные коды [Sinibaldi, Eckman, 2015].

Данные о контактах: дата, время, количество, диспозиционный код

Тип параданных, которые указывают на количество звонков или визитов в определенное домохозяйство, называют в англоязычной литературе «level-of-effort paradata», то есть параданными, указывающими на затраченные интервьюерами усилия. Как правило, фиксируются дата, время и количество контактов или визитов интервьюера, а также диспозиционный код каждого контакта. Такие данные можно использовать как в личных, так и в телефонных опросах. Зачастую они применяются вместе с наблюдениями и оценками интервьюеров.

При анализе параданных могут ставиться следующие задачи:

- стратификация категорий респондентов в зависимости от вероятности участия в опросе и установление приоритета для определенных категорий (с более или менее высокой вероятностью участия — в зависимости от решения исследователей);
- оценка оптимального дня и времени контакта для увеличения вероятности успешно провести интервью;
- оптимизация работы интервьюеров: увеличение количества «эффективных» часов, перераспределение работы между интервьюерами на основе имеющихся данных о загруженности интервьюеров;

- общий расчет необходимого интервьюерам времени, среднего количества контактов на одно завершённое интервью и т. п.;
- оценка ошибки неответов исследования;
- оценка ошибки измерения в зависимости от затраченных усилий на получение интервью.

Данные о датах, времени и количестве контактов позволяют оптимизировать работу интервьюеров. Одно из возможных направлений — предсказание оптимального дня недели и времени визита или звонка. Или, наоборот, предсказание дней недели и времени суток, когда лучше воздержаться от визитов или звонков в домохозяйство определенного типа [Biemer, Chen, Wang, 2013a; Wagner, 2013b]. Другое возможное направление — установление приоритета определенным категориям респондентов для увеличения отклика.

Анализ параданных телефонных опросов в рамках первого подхода (ожидаемо) показал, что доля дозвона выше в выходные, а также в вечерние и утренние часы в будние дни [Durrant, D'Arrigo, Steele, 2011; Mohl, Laflamme, 2007; Purdon, Campanelli, Sturgis, 1999; Weeks, Kulka, Pierson, 1987]. При этом уровень кооперации значительно ниже в вечернее время как в будние, так и в выходные дни, и выше в начале недели (понедельник–вторник), чем в середине и конце недели. Правда, вечером можно с большей вероятностью договориться об интервью в другое время [Durrant, D'Arrigo, Steele, 2013]. Г. Дюррант с коллегами проанализировали параданные шести британских исследований и пришли к выводу, что оптимальный день недели и оптимальное время контакта зависят также и от характеристик домохозяйства. Так, при прочих равных условиях, более высокая вероятность контакта может быть в домохозяйствах с детьми: их можно застать дома как днем, так и вечером. В то время как в домохозяйствах без детей наиболее эффективное время — вечер в будние дни. Кроме того, вероятность следующего дозвона или участия домохозяйства в исследовании зависит также от дня недели и времени предыдущего дозвона. Например, если предыдущий контакт был установлен в будние дни в вечернее время, вероятность того, что интервью может состояться, резко снижается. В этом случае рекомендуется продолжать звонить в вечернее время. Если предыдущий дозвон был в выходные, более эффективной будет попытка установить контакт в ближайшие будние дни в утренние или вечерние часы или в следующие выходные [Durrant, D'Arrigo, Steele, 2011].

В рамках второго подхода исследователи пытаются предсказать наиболее эффективную стратегию для того или иного домохозяйства во время полевых работ. В некоторых ранних работах строились статистические модели, но их эффективность не проверялась в условиях реального исследования (см., например [Greenberg, Stokes, 1990]). В последнее время с увеличением использования параданных в адаптивном дизайне появилась возможность проверить разные стратегии непосредственно в исследованиях. Самой сложной задачей считается оптимизация графика контактов в исследованиях, которые проводятся методом личного интервью дома у респондентов, в особенности при почасовой оплате интервьюеров. С одной стороны, за один приезд в определенный район города интервьюеру необходимо зайти в большее количество домохозяйств, с другой стороны, следует учитывать и рекомендации по опти-

мальному дню и времени визита [Wagner, 2013b]. Эту задачу непросто решать и для телефонных интервью.

Нам пока известно одно опубликованное исследование, в котором вычислялись оптимальные отрезки дня и времени для последующего контакта с тем или иным домохозяйством на основе имеющихся параданных и была предпринята попытка скорректировать работу интервьюеров в ходе полевых работ. Дж. Вагнер провел пять экспериментов в двух американских исследованиях: четыре эксперимента были проведены в опросе методом телефонного интервью и один эксперимент — в опросе методом личного интервью [Wagner, 2013a]. На основе параданных исследователь вычислял оптимальный отрезок времени для звонка или визита (например, рекомендуется позвонить или прийти в домохозяйство в понедельник-четверг с 9.00 до 16.00). В первом эксперименте доля домохозяйств, в которые удалось дозвониться, увеличилась с 9,9% в контрольной группе до 12,0% в экспериментальной группе, однако модель имела нежелательные последствия для тех интервью, в которых ранее был получен отказ. В трех последующих экспериментах в рамках телефонного опроса была предпринята попытка разработать модель с учетом тех, кто ранее отказался принять участие в исследовании, однако она оказалась безуспешной. Наконец, в рамках пятого эксперимента в исследовании методом личного опроса Дж. Вагнер обнаружил, что интервьюеры не следуют инструкциям: доля звонков, сделанных в рекомендуемое время, в контрольной группе составила 23,0%, а в экспериментальной — 23,6%. После завершения полевых работ интервьюеры рассказали, что следовать инструкциям по оптимальному дню недели и времени визита невозможно, так как в случае приезда в определенный район города интервьюер пытается обойти максимальное количество адресов, находящихся в этом районе [Wagner, 2013a]. В целом результаты экспериментов оказались не совсем успешными. Качество статистических моделей, предсказывающих оптимальное время дозвона, было достаточно низким, а интервьюеры не всегда следовали инструкциям исследователей даже в телефонном интервью. В данном случае эти проблемы взаимосвязаны, так как модель строится на основе имеющихся данных о предыдущих контактах. В тех случаях, когда интервьюеры не следуют рекомендациям исследователя, улучшить имеющуюся модель не представляется возможным.

Другое направление использования данных о предыдущих контактах — установление приоритета некоторым категориям респондентов для увеличения их отклика. В рамках адаптивного дизайна это один из самых популярных методов, см. [Groves, Heeringa, 2006; Peytchev et al., 2010; Wagner et al., 2012]. Как правило, приоритет отдается категориям с низкой вероятностью участия. Например, в 2006—2010 гг. в рамках американского исследования «National Survey of Family Growth» Дж. Вагнер с коллегами провели 16 разных экспериментов, направленных на увеличение доли откликов определенных категорий респондентов. На основе имеющихся данных о первичной единице отбора и параданных (о предыдущих контактах и наблюдениях интервьюеров) в разных фазах процесса сбора данных исследователи устанавливали приоритет определенным категориям домохозяйств [Wagner et al., 2012]. На основе анализа данных о последующих контактах исследователи приходят к выводу, что интервьюеры старались следовать инструкциям,

и в семи из шестнадцати экспериментов количество визитов в эти домохозяйства значительно увеличилось. Однако доля откликов среди этих домохозяйств выросла только в двух экспериментах [Wagner et al., 2012].

Недостатки и сложности использования параданных в адаптивном дизайне

Выше мы фокусировали внимание на преимуществах и возможностях использования параданных в адаптивном дизайне. Однако необходимо отметить и сложности при их получении и использовании.

С наблюдениями интервьюеров сопряжены пять основных сложностей. (1) Точность наблюдений может быть невысокой и варьироваться от одной переменной к другой. (2) Эти наблюдения и оценки подвержены систематической ошибке измерения и зависят как от характеристик самого интервьюера, так и от характеристик домохозяйства и состава района проживания. Например, было показано, что в районах со значительной долей афроамериканских или латиноамериканских домохозяйств точность оценок интервьюеров снижается [West, Kreuter, 2013]. (3) Точность ответов и понимание вопросов интервьюерами может зависеть от формулировки вопросов, которые исследователи задают интервьюерам для записи наблюдений [Kaminska, Lynn, 2011]. (4) Ведение наблюдений — задача, которая дополнительно ложится на плечи интервьюеров, что может сказаться как на качестве наблюдений, так и на качестве проведения основного опроса. (5) Сбор дополнительных данных, как правило, увеличивает общий бюджет исследования [Sinibaldi, Trappmann, Kreuter, 2014].

В личных интервью трудности могут возникнуть при сборе данных о контактах. Если в телефонных интервью такие данные почти полностью собираются автоматически, то в личных опросах интервьюеры должны фиксировать дату и время каждого визита в домохозяйство. Можно предположить, что интервьюеры не всегда это делают в полном объеме. Так, в одном американском исследовании на основе качества статистической модели исследователи предположили, что интервьюеры фиксируют данные не обо всех контактах, что подтвердили и сами интервьюеры [Biemer, Chen, Wang, 2013]. Кроме того, ведение бланка контактов увеличивает общий бюджет исследования в опросах методом личного интервью.

В целом сила моделей, предсказывающих вероятность участия домохозяйства на основе имеющихся наблюдений интервьюеров и данных о предыдущих контактах, может быть достаточно невысокой. Например, использование только данных с основным описанием первичной единицы отбора (город/село, наличие этнического меньшинства в регионе, наличие правительственных органов в регионе) и наблюдений интервьюеров на данных панельного исследования «Understanding Society» может объяснить 6% дисперсии, когда предсказывается количество звонков, необходимое для получения интервью, и 5% дисперсии, когда предсказывается, будет ли проведено интервью или нет. Если модель дополнить информацией о первом контакте, то доля объясненной дисперсии увеличивается до 8% и 6,5%, соответственно. Если включить информацию по двум контактам, то доля объясненной дисперсии увеличивается до 12% и 10%, соответственно [Durrant, Maslovskaya, Smith, 2015].

Другие исследования также показали, что доля объясненной дисперсии, как правило, не превышает 10—15 % в случае предсказания завершеного интервью [Sinibaldi, Eckman, 2015; Sinibaldi, Trappmann, Kreuter, 2014; Sturgis, Brunton-Smith, 2011]. Дополнительную сложность в работе с параданными в рамках адаптивного дизайна представляет тот факт, что исследователи ни в какой момент времени не имеют полных данных [Wagner, Hubbard, 2014]. Например, Г. Дюррант с коллегами показали, что добавление информации о третьем контакте увеличивает процент объясненной дисперсии с 10 до 24 % в предсказании участия домохозяйства в опросе [Durrant, Maslovskaya, Smith, 2015]. А. Блом демонстрирует зависимость предсказательной силы моделей от страны: так, например, на данных Европейского социального исследования процент объясненной дисперсии в моделях, предсказывающих установление контакта с домохозяйством в случае визита интервьюера, варьировался от 25—26 % в Словакии и Польше до 48 % в Финляндии. Включение дополнительных переменных, таких как информация об интервьюере и домохозяйстве (например, размер домохозяйства), может увеличить объясненную дисперсию до 46—58 % в попытках предсказать вероятность получения интервью [Blom, 2009].

Дополнительная проблема для адаптивного дизайна, которая может внести смещения, состоит в том, что респонденты, участвующие в исследовании в начале, отличаются от респондентов, которых удастся опросить ближе к завершению полевых работ [Hox, de Leeuw, Chang, 2012; Kaminska, Mccutcheon, Billiet, 2010; Triplett, Blair, Kang, 1996]. Таким образом, анализ параданных и внесение изменений в дизайн на основе имеющихся в начале полевых работ данных может не привести ни к снижению ошибки неответов, ни к увеличению доли откликов. В этом слабость адаптивного дизайна.

Заключение

В последние годы растет число публикаций, в которых обсуждается адаптивный дизайн и возможности использования в нем параданных. Исследователи включают сюда наблюдения интервьюеров; данные о количестве визитов или звонков в домохозяйство, а также дату, время и диспозиционный код каждого визита или звонка; время заполнения анкеты и паттерны заполнения анкеты (лог файлы, включающие клики мышки, нажатие клавиш, скорость печати и т. д.); GPS-маршруты интервьюеров или GPS-координаты определенных интервью [Olson, Wagner, 2015] и т. п. В обзоре мы сфокусировали внимание только на наблюдениях интервьюеров и данных о контактах (дата, время, количество, диспозиционные коды), которые могут использоваться для внесения изменений в дизайн исследования в ходе полевых работ.

Мы рассмотрели отдельно наблюдения интервьюеров и данные о контактах, однако подобное разделение искусственно, если мы говорим о реальных практиках их применения в адаптивном дизайне. Как правило, оба вида параданных используются в моделях одновременно. Ясно, что исследователю необходимо иметь на руках как можно больше данных для корректировки и оптимизации полевых работ. Следует учитывать и разные типы параданных, и информацию о первичных единицах отбора, и, возможно, любые дополнительные данные, пригодные для уточнения моделей.

В настоящее время пока нет понимания, какие стратегии более эффективны, какие элементы дизайна лучше менять исследователю, а также к каким дополнительным смещениям они могут привести. Мы показали, что некоторые эксперименты в рамках адаптивного дизайна были успешными, в то время как другие не привели к желаемому результату. Можно предположить, что с увеличением количества экспериментальных данных мы получим больше возможностей для построения эффективного адаптивного дизайна, ключевую роль в котором будет играть сбор и анализ параданных.

Сделанный обзор не полон и показывает лишь небольшую часть того, как можно использовать параданные в опросах и, в частности, в рамках адаптивного дизайна.

Список литературы (References)

Ипатова А. А. Опыт анализа комментариев интервьюеров к телефонным интервью // Мониторинг общественного мнения: Экономические и социальные перемены. 2016. № 5. С. 64—76. DOI: 10.14515/monitoring.2016.5.05.

Ipatova A. A. (2016) The experience in the analysis of the interviewers comments to telephone interviews. *Monitoring of Public Opinion: Economic and Social Changes*. No. 5. P. 64—76. (In Russ.)

Ипатова А. А. Использование параданных в анализе телефонных опросов // Телескоп. 2014. № 6 (108). С. 34—41.

Ipatova A. A. (2014) Using paradata in the analysis of telephone surveys. *Teleskop [Telescope]*. No. 6 (108). P. 34—41. (In Russ.)

Рогозин Д. М., Ипатова А. А. Контроль качества социальных обследований. М., 2015. URL: <ftp://w82.ranepa.ru/rnp/ppaper/020915.pdf> (дата обращения: 13.08.2017).

Rogozin D. M., Ipatova A. A. (2015) Control over the quality of social surveys. Moscow. Available at: <ftp://w82.ranepa.ru/rnp/ppaper/020915.pdf> (accessed: 13.08.2017). (In Russ.)

Рогозин Д. М., Ипатова А. А. Как быстро, эффективно и безопасно сфабриковать телефонный опрос // Социологический журнал. 2012. № 2. С. 80—97.

Rogozin D. M., Ipatova A. A. (2012) How to quickly, effectively and safely fabricate a telephone survey. *Sociological Journal*. No. 2. P. 80—97. (In Russ.)

Сидоров Н. Использование параданных в социологических исследованиях // Социология: теория, методы, маркетинг. 2011. 4. С. 198—208.

Sidorov N. (2011) Using paradata in sociological surveys. *Sociology: Theory, Methods, Marketing*. 4. P. 198—208. (In Russ.)

Axinn W. G., Link C., Groves R. M. (2011) Responsive survey design, demographic data collection, and models of demographic behavior. *Demography*. Vol. 48. No. 3. P. 1127—1149. DOI: 10.1007/s13524-011-0044-1.

Biemer P., Chen P., Wang K. (2013) Using level-of-effort paradata in non-response adjustments with application to field surveys. *Journal of the Royal Statistical Society: Series A*. Vol. 176. No. 1. P. 147—168. DOI: 10.1111/j.1467-985X.2012.01058.x.

Blom A. G. (2009). Non-response Bias Adjustments: What Can Process Data Contribute? Institute for Social and Economic Research Working Paper Series. No. 2009—21.

Couper M. P. (1998) Measuring survey quality in a CASIC environment. *Proceedings of the Survey Research Methods Section. American Statistical Association*. P. 41—49.

Couper M. P. (2005) Technology trends in survey data collection. *Social Science Computer Review*. Vol. 23. No. 4. P. 486—501. DOI: 10.1177/0894439305278972.

Durrant G. B., D'Arrigo J., Steele F. (2011) Using paradata to predict best times of contact, conditioning on household and interviewer influences. *Journal of the Royal Statistical Society: Series A*. Vol. 174. No. 4. P. 1029—1049. DOI: 10.1111/j.1467—985X.2011.00715.x

Durrant G. B., D'Arrigo J., Steele F. (2013) Analysing interviewer call record data by using a multilevel discrete-time event history modelling approach. *Journal of the Royal Statistical Society: Series A*. Vol. 176. No. 1. P. 251—269. DOI: 10.1111/j.1467-985X.2012.01073.x.

Durrant G., Maslovskaya O., Smith P. W. F. (2015). Modeling final outcome and length of call sequence to improve efficiency in interviewer call scheduling. *Journal of Survey Statistics and Methodology*. Vol. 3. No. 3. P. 397—424. DOI: 10.1093/jssam/smv008.

Greenberg B. S., Stokes S. L. (1990). Developing an optimal call scheduling strategy for a telephone survey. *Journal of Official Statistics*. Vol. 6. No. 4. P. 421—435.

Groves R. M., Heeringa S. G. (2006) Responsive design for household surveys: tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society: Series A*. Vol. 169. No. 3. P. 439—457. DOI: 10.1111/j.1467-985X.2006.00423.x.

Eckman S., Sinibaldi J., Möntmann-Hertz A. (2013) Can interviewers rate the likelihood of cases to cooperate? *Public Opinion Quarterly*. Vol. 77. No. 2. P. 561—573. DOI: 10.1093/poq/nft012.

Hox J., de Leeuw E., Chang H.-T. (2012). Nonresponse versus measurement error: Are reluctant respondents worth pursuing? *Bulletin de Methodologie Sociologique*. Vol. 113. No. 1. P. 5—19. DOI: 10.1177/07591063111426987.

Kaminska O., Lynn P. (2011) Interviewer observation: are the observations what we want or what interviewers think we want? Paper presented at Nonresponse Workshop. Bilbao, Spain.

Kaminska O., Mccutcheon A. L., Billiet J. (2010). Satisficing among reluctant respondents in a cross-national context. *Public Opinion Quarterly*. Vol. 74. No. 5. P. 956—984. DOI: 10.1093/poq/nfq062.

Kreuter F. (2013) Improving surveys with paradata: Introduction. In: Kreuter F. Improving surveys with paradata. Analytic Uses of Process Information. New Jersey: John Wiley and Sons. P. 1—9.

Kirgis N. G., Lepkowski J. M. (2013) Design and management strategies for paradata-driven responsive design: Illustrations from the 2006—2010 National Survey of Family Growth. In: Kreuter F. Improving surveys with paradata. Analytic Uses of Process Information. New Jersey: John Wiley and Sons. P. 123—144.

Mohl C., Laflamme F. (2007) Research and responsive design options for survey data collection at statistics Canada. *Proceedings of the Survey Research Methods Section, Joint Statistical Meetings. Section on Survey Research Methods*. P. 2962—2968. URL: <https://ww2.amstat.org/sections/srms/Proceedings/y2007/Files/JSM2007-000421.pdf> (accessed 26.10.2017).

Olson K., Wagner J. (2015) A feasibility test of using smartphones to collect GPS information in face-to-face surveys. *Survey Research Methods*. Vol. 9. No. 1. P. 1—13. DOI: 10.18148/srm/2015.v9i1.6036.

Purdon S., Campanelli P., Sturgis P. (1999) Interviewers' calling strategies on face-to-face interview surveys. *Journal of Official Statistics*. Vol. 15. No. 2. P. 199—216.

Peytchev A., Riley S., Rosen J., Murphy J., Lindblad M. (2010) Reduction of nonresponse bias in surveys through case prioritization. *Survey Research Methods*. Vol. 4. No. 1. P. 21—29. DOI: 10.18148/srm/2010.v4i1.3037.

Sinibaldi J., Eckman S. (2015). Using call-level interviewer observations to improve response propensity models. *Public Opinion Quarterly*. Vol. 79. No. 4. P. 976—993. DOI: 10.1093/poq/nfv035.

Sinibaldi J., Trappmann M., Kreuter F. (2014) Which is the better investment for nonresponse adjustment: Purchasing commercial auxiliary data or collecting interviewer observations? *Public Opinion Quarterly*. Vol. 78. No. 2. P. 440—73. DOI: 10.1093/poq/nfu003.

Schouten B., Calinescu M., Luiten A. (2013) Optimizing quality of response through adaptive survey designs. *Survey Methodology*. Vol. 39. No. 1. P. 29—58.

Schouten B., Peytchev A., Wagner J. (2017) Adaptive survey design. Florida, CRC Press.

Sturgis P., Brunton-Smith I. (2011) An assessment of the potential utility of interviewer observation variables for reducing non-response error in the National Survey for Wales. A report prepared for the Welsh Government. <https://eprints.soton.ac.uk/209597/1/FINALREPORTjan12.pdf> (accessed 15.08.2017).

Tourangeau R., Brick J. M., Lohr S., Li J. (2017) Adaptive and responsive survey designs: A review and assessment. *Journal of the Royal Statistical Society: Series A*. Vol. 180. No. 1. P. 203—223. DOI: 10.1111/rssa.12186.

- Triplett T., Blair, J., Kang, Y.C. (1996). Initial cooperators vs. converted refusers: Are there response behavior differences? *Proceedings of the Survey Research Methods Section. American Statistical Association*. P. 1038—1041.
- Wagner J. (2008) Adaptive survey design to reduce nonresponse bias. A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Survey Methodology) in the University of Michigan. URL: https://deepblue.lib.umich.edu/bitstream/handle/2027.42/60831/jameswag_1.pdf (accessed 15.08.2017).
- Wagner J. (2013a) Adaptive contact strategies in telephone and face-to-face surveys. *Survey Research Methods*. Vol. 7. No. 1. P. 45—55. DOI: 10.18148/srm/2013.v7i1.5037.
- Wagner J. (2013b.) Using paradata-driven models to improve contact rates in telephone and face-to-face surveys. In: Kreuter F. Improving surveys with paradata. Analytic Uses of Process Information. New Jersey: John Wiley and Sons. P. 145—170.
- Wagner J., Hubbard F. (2014) Producing unbiased estimates of propensity models during data collection. *Journal of Survey Statistics and Methodology*. Vol. 2. No. 3. 323—342. DOI: doi.org/10.1093/jssam/smu009.
- Wagner J., West B., Kirgis N., Lepkowski J. M., Axinn W. G., Kruger N. S. (2012). Use of paradata in a responsive design framework to manage a field data collection. *Journal of Official Statistics*. Vol. 28. No. 4. P. 477—499.
- Weeks M. F., Kulka R. A., Pierson S. A. (1987) Optimal call scheduling for a telephone survey. *Public Opinion Quarterly*. Vol. 51. No. 4. P. 540—549. DOI: 10.1086/269056.
- West B. (2013) An examination of the quality and utility of interviewer observations in the National Survey of Family Growth. *Journal of the Royal Statistical Society: Series A*. Vol. 176. No. 1. P. 211—225. DOI: 10.1111/j.1467—985X.2012.01038.x.
- West B., Kreuter F. (2013) Factors affecting the accuracy of interviewer observations: Evidence from the National Survey of Family Growth. *Public Opinion Quarterly*. Vol. 77, No. 2, pp. 522—548. DOI: 10.1093/poq/nft016.