

ТЕОРИЯ И МЕТОДОЛОГИЯ



DOI: 10.19181/socjour.2026.32.3.1
EDN: BQYDRW

Е.Д. ШМЕЛЕВА¹, Д.В. ЛЕБЕДЕВ¹, Е.В. БРАГИНА¹

¹ Национальный исследовательский университет «Высшая школа экономики». 109028, Москва, Покровский бульвар, д. 11.

«А ВСЕ-ТАКИ...»: ЭФФЕКТИВНОСТЬ УТОЧНЯЮЩИХ ВОПРОСОВ ДЛЯ СНИЖЕНИЯ ДОЛИ НЕОТВЕТОВ НА ВОПРОСЫ О ФИНАНСАХ¹

Аннотация. Высокая доля несодержательных ответов на вопросы о доходах и расходах снижает качество данных, ограничивает генерализацию результатов и может приводить к смещениям в выборке. Цель исследования — оценить эффективность методики уточняющих вопросов для повышения качества ответов о финансах. Эмпирическая основа — данные 6-й волны проекта «Экономическое поведение домашних хозяйств» НИУ ВШЭ (весна 2025, N = 6 041). На первом этапе определялись предикторы несодержательных ответов, на втором оценивалась эффективность уточняющих вопросов для разных категорий респондентов. Отбор предикторов проводился с помощью метода случайного леса (Random Forest) с алгоритмом Boruta и логистических регрессий. Результаты показывают, что доля несодержательных ответов составила 6% для индивидуальных доходов, 17% для семейных доходов, 14% для семейных расходов. Уточняющие вопросы преобразовали в содержательные 36–41% первоначальных неответов, то есть сократили общий уровень неответов до 4–10%. Наибольший риск неответа выявлен у респондентов, живущих отдельно с супругом (рост шансов более чем в 2 раза), неосведомленных о семейных расходах (в 3–5 раз), с депрессивными состояниями (на 18–32%), плохим здоровьем (на 50–86%), без сбережений, с коротким горизонтом планирования и жителей небольших населенных пунктов. Уточняющие вопросы наиболее эффективны для респондентов с низким образованием, неосведомленных о расходах, без сбережений, из малых городов. Методика эффективна лишь частично: она снижает когнитивную сложность, но не влияет на чувствительность вопросов о доходах. На основе результатов предложены рекомендации по улучшению дизайна опросов, включающих финансовые вопросы.

Ключевые слова: качество опросных данных; несодержательные ответы; проблема неответа; ответ «затрудняюсь ответить»; чувствительные вопросы; уточняющие вопросы; вопросы о финансах; доходы; расходы.

¹ В данной работе использованы результаты проекта HSE-MS-2025-006, выполненного в НИУ ВШЭ.

Для цитирования: Шмелева Е.Д., Лебедев Д.В., Брагина Е.В. «А все-таки...»: эффективность уточняющих вопросов для снижения доли неответов на вопросы о финансах // Социологический журнал. 2026. Том 32. № 3. С. 9–29. DOI: [10.19181/socjour.2026.32.3.1EDN: BQYDRW](https://doi.org/10.19181/socjour.2026.32.3.1EDN:BQYDRW)

Введение

Вопросы о доходах и расходах относятся к числу вызывающих наибольшие затруднения у респондентов [25; 3]. На это указывает высокая доля несодержательных ответов — «затрудняюсь ответить» и «отказ от ответа»: анализ девяти исследований, проведенных в США с 1999 по 2005 г., показал, что доля отказывающихся от ответа на вопросы о доходах варьируется от 14 до 35% [30]². Данные Европейского социального исследования (ESS) демонстрируют аналогичную тенденцию: доля избегающих ответов на финансовые вопросы превышает показатели по другим темам и достигает 20–40% в Чехии, Венгрии, Ирландии, Польше, Швейцарии и Испании [13].

В постсоветских странах, включая Россию, пропуски данных о доходах достигают 18–45% (в среднем 30%). Исследователи связывают это с негативными эмоциями, которые финансовые вопросы вызывают у респондентов, — многие испытывают «враждебность и возмущение» [4, с. 24].

Проблема неответов на финансовые вопросы серьезно влияет на качество результатов исследований, даже если основная тема опросов не связана с финансами. Данные о доходах и расходах широко используются в социологических, государственных, маркетинговых исследованиях, где часто служат основой для формирования выборки и обеспечения репрезентативности опросов населения [1; 3]. Высокая доля неответов на эти вопросы снижает точность социально-экономических измерений и препятствует корректной оценке взаимосвязей между переменными, которая может быть искажена из-за смещений в выборке [13; 5]. На уровне государственной статистики США, например, обнаруживаются систематические смещения в ответах про доходы, что приводит к завышению статистики доходов на 2–3% и занижению официальных показателей бедности³. Вопросы о доходах и расходах населения критически важны для разработки программ пенсионного обеспечения и оценки кредитных рисков, так как позволяют прогнозировать устойчивость экономики к финансовым шокам на уровне отдельных домохозяйств [21].

Для снижения проблемы неответов в вопросах о финансах исследователи применяют различные методики, которые призваны снизить когнитивную нагрузку и чувствительность реакции на вопросы — основные факторы выбора несодержательных ответов [13]. К числу таких методик относятся уточняющие вопросы (“probing”), среди которых выделяются метод «раскрывающихся скобок» (unfolding brackets) [12; 18], упрощенная версия с заданными интервалами, среди которых можно выбрать диапазон значений доходов/расходов [20], а также детализированные вопросы о каждом из компонентов доходов [8; 15].

² Более подробный обзор см. в книге А.Ю. Мягкова «Косвенные методы в сенситивных исследованиях» [3].

³ Using administrative data to assess nonresponse bias in the Current Population Survey 2023. Annual Social and Economic Supplement. U.S. Census Bureau. — URL: <https://www.census.gov/newsroom/blogs/research-matters/2023/09/using-administrative-data-nonresponse-cps-asec.html> (дата обращения 20.10.2025).

Цель данного исследования состоит в оценке эффективности методики уточняющего вопроса с заданными интервалами для определения уровня доходов и расходов (семейных и индивидуальных) на основе данных выборочного опроса 6 волны проекта «Экономическое поведение домашних хозяйств», охватывающего население России в возрасте 18 лет и старше и проведенного НИУ ВШЭ весной 2025 г. Для этого на основе обзора литературы и разведывательного метода анализа Random Forest (с применением метода оценки важности переменных Boruta [19]) мы определяем потенциальные предикторы несодержательных ответов на три открытых вопроса: 1) об индивидуальном доходе, 2) о семейных доходах, а также 3) о семейных расходах. Далее с помощью серии логистических регрессий определяем наиболее значимые предикторы выбора несодержательных ответов на вопросы о финансах. Затем мы оцениваем доли тех, кто в уточняющих вопросах, задаваемых после открытых вопросов о доходах/расходах, дает содержательный ответ, то есть выбирает какой-то из предложенных диапазонов расходов/доходов. Это позволяет нам определить, для каких категорий респондентов уточняющие вопросы повышают шансы дать содержательный ответ, а для каких — нет.

Обзор литературы

Факторы высокой доли несодержательных ответов на вопросы о финансах

Вопросы о финансовом положении респондентов традиционно относятся к наиболее проблематичным [3], поскольку демонстрируют высокие доли несодержательных ответов («затрудняюсь ответить», «отказ от ответа», пропуски вопросов), что существенно снижает качество данных и может приводить к систематическим смещениям в выборке.

Исследователи выделяют два главных фактора, объясняющих высокую проблематичность финансовых вопросов: их когнитивную сложность и сенситивность (чувствительность) [14].

Первая связана с объективными трудностями формулирования ответа в ходе двух когнитивных процессов — проблем с припоминанием и согласованием информации [11]. Респонденты могут не знать точной информации, сталкиваться с необходимостью идентифицировать множественные источники доходов и расходов и с трудностями производить сложные вычисления [25]. Особенно проблематичными являются вопросы относительно общей стоимости активов, инвестиций или наследства, требующие детальных подсчетов и финансовой грамотности, а также вопросы о финансах других членов домохозяйства, требующие агрегирования информации и осведомленности [13]. При этом, согласно принципу «разумной достаточности» или «минимально достаточных усилий» (satisficing) [17], когда когнитивная нагрузка превышает заинтересованность респондента в качественном заполнении анкеты, он выбирает стратегию минимизации усилий, отдавая предпочтение несодержательным ответам. По этой причине неответы на подобные вопросы могут быть следствием незаинтересованности или когнитивной «усталости» респондента и нежеланием затрачивать когнитивные усилия на предоставление точного ответа.

Сенситивность финансовых вопросов обусловлена их восприятием как слишком личных или потенциально угрожающих [28]. В методологии опросов к сенситивным относятся вопросы, затрагивающие сферу частной жизни, способные повлечь социальные последствия или связанные с социальным сравнением [28]. Для финансовых вопросов это выражается в опасениях продемонстрировать «недостаточное» материальное благополучие, подвергнуться осуждению или в нежелании раскрывать информацию, которая считается сугубо частной, особенно в индивидуалистических обществах, где ценится приватность личной финансовой информации [13].

Факторы выбора несодержательных ответов

Склонность к выбору несодержательных ответов зависит от социально-демографических характеристик респондентов, их когнитивных способностей, культурного контекста и параметров дизайна опроса⁴. К социально-демографическим характеристикам, наиболее часто встречающимся в исследованиях, относятся возраст, пол, образование, уровень дохода и занятость, размер домохозяйства и место жительства. Пожилые респонденты более склонны к отказам отвечать на вопросы о доходах [9; 25] и чаще изначально выбирают вариант «не знаю» [18]. Женщины чаще мужчин выбирают несодержательный ответ [30]. Риск неответов также выше для респондентов, состоящих в браке и проживающих в больших домохозяйствах [13; 14; 25].

Респонденты с более высоким уровнем образования реже выбирают несодержательные ответы о доходах, предположительно из-за меньших когнитивных трудностей, вызываемых финансовыми вопросами [14; 30; 16]. Аналогичная закономерность может касаться и состояния здоровья: респонденты с худшим самочувствием располагают меньшими когнитивными ресурсами для обработки сложных вопросов, что также повышает вероятность несодержательного ответа. Оба эффекта зафиксированы в крупном популяционном исследовании (N = 360 628), где была установлена отрицательная взаимосвязь между склонностью затрудняться с ответом и уровнем образования, а также самооценкой здоровья [22]. Вместе с тем ряд исследований фиксируют противоположные результаты, судя по которым образованные респонденты чаще отказываются отвечать на вопросы о доходах, воспринимая их как слишком личные, то есть из соображений сенситивности, а не из-за когнитивных затруднений [24].

Склонность к выбору несодержательных ответов на вопросы о доходах варьируется в зависимости от социально-экономического положения респондента: люди с низкими доходами чаще выбирают прямой отказ от ответа — преимущественно из-за сенситивности темы, — тогда как у высокодоходных групп одинаково вероятны оба типа неответов, что свидетельствует о дополнительных трудностях с припоминанием и агрегированием финансовой информации [14]. Самозанятые чаще склонны к отказу от ответа, тогда как работники госсектора, напротив, демонстрируют большую готовность делиться этой информацией [24]. В целом риск неответов возрастает при нестабильных доходах и при наличии дополнительных источников дохода, например инвестиционных [13; 14].

⁴ Имеются в виду прежде всего условия сбора данных, такие как пол, возраст интервьюера и их совпадение с этими характеристиками респондентов [24; 25].

Методики снижения риска неответов на финансовые вопросы

Для снижения доли несодержательных ответов исследователи применяют различные методики в рамках дизайна опросника. Например, для снижения когнитивной нагрузки при агрегировании доходов из различных источников применяются детализированные вопросы о каждом из компонентов дохода [8; 15]. Большую популярность приобрели уточняющие вопросы, три основных типа которых описали Л. Шварц и Г. Паулин [27]. «Стандартная» (*conventional*) методика предлагает готовые интервалы доходов в уточняющем вопросе после открытого, «разворачивающаяся» (*unfolding*) предполагает последовательные вопросы для постепенного сужения диапазона⁵, а «метод интервалов, генерируемых респондентами» (*respondent-generated intervals — RGI*), позволяет участникам самостоятельно определить минимальные и максимальные границы своих доходов. Стандартную методику охотнее выбирают более пожилые люди и с невысоким уровнем образования. Метод RGI предпочитают люди с высшим образованием, так как он дает им больше контроля, — и часто такой подход приводит к более точным результатам [27].

В целом методика уточняющих вопросов демонстрирует эффективность в снижении доли неответов на нефинансовые вопросы. В эксперименте, где сравнивался отклик на псевдознаниевые вопросы⁶ и вопросы об установках⁷, в половине которых (случайной) использовались последующие уточняющие вопросы, а в другой — нет, почти половина изначально затруднившихся потом ответили на уточняющий вопрос [18].

В отношении финансов методика уточняющих вопросов позволяет сократить долю пропущенных данных до 56% [30]. Эффективность этой методики, однако, ограничена: если респондент уклоняется от ответа из-за опасений по поводу конфиденциальности (проблема сенситивности), вопрос с заданными интервалами значений дохода все равно может быть отвергнут [29].

Эмпирическая база

Эмпирической основой исследования являются опросные данные 6-й волны опроса «Экономическое поведение домашних хозяйств»⁸, охватывающего население России в возрасте 18 лет и старше, проведенного НИУ ВШЭ весной 2025 г. Опрос проводился методом CAPI (личное интервью с занесением информации в программу через планшет) по многоступенчатой стратифицированной выборке с маршрутным отбором домохозяйства для обследования и опросом одного пред-

⁵ Например: «За последние 12 месяцев получали ли вы заработную плату? Ваша заработная плата более 3 млн рублей?» Если да, то: «Ваша заработная плата более 3,2 млн рублей?»; Если нет, то: «Ваша заработная плата более/менее 2,8 млн рублей?» и далее подобные развилки вопросов.

⁶ Например: «Из каждых 100 человек трудоспособного возраста в [стране] сколько, по вашему мнению... безработные и ищут работу?» Варианты ответа: 0–4 (1), 5–9 (2), 10–14 (3), 15–19 (4), 20–24 (5), 25–29 (6), 30–34 (7), 35–39 (8), 40–44 (9), 45–49 (10), 50 и более (11).

⁷ Например: «Пожалуйста, скажите, насколько вы согласны или не согласны с каждым из следующих утверждений о людях в [стране]. Большинство безработных на самом деле не пытаются найти работу». Варианты ответа: от «Совершенно не согласен(-на)» до «Полностью согласен(-на)».

⁸ Более подробно об исследовании на сайте проекта: <https://www.hse.ru/hheb/>

ставителя внутри домохозяйства. Объем выборки составил 6 041 чел., выборка репрезентирует все население РФ по федеральным округам, типу населенного пункта, полу и возрасту.

Аналитическая стратегия

Исследование реализовано в два этапа, отражающих две ключевые задачи: 1) определить предикторы несодержательных ответов на открытые вопросы о доходах и расходах, а также 2) определить предикторы выбора содержательных ответов в уточняющих закрытых вопросах с интервалами значений. Для каждого этапа применяется двухэтапный подход, начинающийся с использования алгоритма Random Forest для первичного отбора наиболее важных предикторов среди переменных методом Boruta [19], выделенных на основе обзора литературы, с последующим применением серии логистических регрессий для точного определения статистически значимых предикторов (на уровне $\alpha = 0,05$) и оценки направления взаимосвязи между ними.

Для отбора ключевых переменных использовался метод Random Forest с алгоритмом Boruta. Метод Random Forest строит набор деревьев решений (Decision Trees) на случайных подвыборках и оценивает, какие переменные лучше предсказывают результат [6]. Алгоритм Boruta позволяет определить наиболее важные переменные (действительно связанные с зависимой переменной), сравнивая их вклад в предсказание значений зависимой переменной с искусственно созданными теневыми копиями, не содержащими полезной информации [19]. Такой подход рекомендуется для выделения важности переменных для последующего анализа [10; 26], так как позволяет выделять сложные и нелинейные связи между переменными и отбирать только действительно важные признаки для последующего анализа [6].

Преимущества выбранного метода для определения ключевых переменных, связанных с выбором несодержательного ответа и валидным ответом на уточняющий вопрос, по сравнению с более традиционными аналитическими методами (например, логистическая регрессия или однофакторный дисперсионный анализ) заключаются в следующем [6; 10; 19]:

1. непараметрический характер метода — отсутствие строгих предпосылок о характере распределения переменных;
2. возможность учитывать нелинейные связи и взаимодействия автоматически;
3. алгоритм Boruta позволяет решать задачу отбора признаков на основе принципа «все релевантные переменные» (all-relevant feature selection), а не только минимального оптимального набора;
4. метод устойчив к шуму и мультиколлинеарности.

Стоит отметить также и наличие некоторых ограничений выбранного метода анализа [6; 10; 24]. Во-первых, интерпретируемость моделей случайного леса ниже по сравнению с традиционными методами. Однако в нашем случае модели случайного леса с применением алгоритма Boruta используются строго для определения набора релевантных характеристик, а не для интерпретации связи с зависимыми переменными. Во-вторых, алгоритм Boruta может быть вычислительно затратным, особенно при большом числе признаков и наблюдений. В нашем случае вычисления занимали в среднем от 30 до 50 минут.

Технически метод Random Forest с применением алгоритма Boruta был реализован с использованием пакета Boruta версии 8.0.0 [19], с размером леса в Random Forest — 500. Вся обработка данных и их анализ были произведены в R Studio (версия 4.4.1)⁹.

Измерения

Зависимые переменные

В качестве зависимых переменных используются ответы на вопросы (табл. 1), измеряющие размер 1) личного дохода, 2) семейных совокупных доходов, 3) семейных совокупных расходов, а также ответы на уточняющие вопросы, где в качестве вариантов ответа выступают интервалы значений размеров доходов и расходов с шагом в 10 000 рублей. Для анализа переменные были преобразованы следующим образом: в исходных вопросах о размере доходов и расходов содержательные ответы были закодированы в 0, а несодержательные («Затрудняюсь ответить», «Отказ от ответа») — в 1; в уточняющих вопросах, наоборот, содержательные ответы на вопросы были закодированы в 1, а несодержательные («Затрудняюсь ответить», «Отказ от ответа») — в 0.

Таблица 1

Вопросы, измеряющие размер индивидуальных, семейных доходов и расходов, а также уточняющие вопросы с диапазонами значений

Концепт	Открытый вопрос	Закрытый уточняющий вопрос
Размер индивидуального дохода (задавался всем работающим, вне зависимости от формата работы)	Каков примерно размер вашего заработка, доходов от основной работы в прошлом месяце, включая премии, отпускные и прочие выплаты (после вычета налогов)? 99. (не зачитывайте) Затрудняюсь ответить, отказ от ответа	А все-таки, к какой из групп на этой карточке можно было бы отнести ваш заработок, доход от основной работы в прошлом месяце, включая премии, отпускные и прочие выплаты (после вычета налогов)? <i>Варианты ответа:</i> 1. Менее 10 000 рублей в месяц 2. 10 001 – 20 000 рублей в месяц 3. 20 001 – 30 000 рублей в месяц <...> 9. 150 001 – 200 000 рублей в месяц 10. Более 200 000 рублей в месяц 99. (не зачитывайте) Затрудняюсь ответить 98. (не зачитывайте) Отказ от ответа

⁹ Код обработки и анализа данных, а также дополнительные материалы (графики по результатам Random Forest, таблицы с результатами логистических регрессий, описание зависимых и независимых переменных и их подготовки к анализу, скрипты R, использованные для обработки и анализа данных) доступны в репозитории по ссылке: <https://disk.360.yandex.ru/d/PUeCaJDaTj2B2w>

Продолжение таблицы 1

Концепт	Открытый вопрос	Закрытый уточняющий вопрос
Размер доходов семьи (задавался всем)	А теперь, учитывая все отмеченные вами источники денежных доходов, подсчитайте, пожалуйста, каким был ваш и всех членов вашей семьи, проживающих вместе с вами, общий (суммарный) доход в прошлом месяце. Если какие-либо доходы были получены в валюте, переведите их в рубли по нынешнему обменному курсу. 99. (не зачитывайте) Затрудняюсь ответить, отказ от ответа	А все-таки, к какой из групп на этой карточке можно было бы отнести доход вашей семьи в прошлом месяце? Варианты ответа: аналогично вопросу выше
Размер расходов домохозяйства (задавался всем)	Какую примерно сумму в рублях вы и ваша семья потратили за последние 30 дней? Пожалуйста, посчитайте все расходы, учитывая платежи за аренду жилья, выплаты по ипотеке, кредиты, коммунальные платежи, пожертвования, а также расходы на покупку еды, одежды, оплату транспорта, развлекательных мероприятий и все другие расходы, которые были в вашей семье. 99. (не зачитывайте) Затрудняюсь ответить, отказ от ответа	А все-таки, к какой категории можно было бы отнести расходы вашей семьи в предыдущем месяце? Варианты ответа: аналогично вопросу выше

Независимые переменные

В качестве независимых используются переменные (интервальные, а также категориальные, преобразованные для анализа в дамми-переменные), отражающие¹⁰:

- социально-демографические характеристики: пол, возраст, размер семьи, семейное положение, размер населенного пункта;
- занятость и уровень образования: образование (7 уровней от начального до наличия степени), занятость (10 статусов от учебы до статуса безработного);
- финансовое поведение и установки: наличие сбережений и долгов; горизонт планирования; уровень финансовой тревожности; осведомленность о финансах других членов семьи; доля расходов на еду; субъективная оценка материального положения; количество автомобилей в семье;

¹⁰ Более подробное описание переменных представлено в репозитории с дополнительными материалами исследования по ссылке: <https://disk.360.yandex.ru/d/PUEcaJDaTj2B2w>

- представления о доходах, необходимых для «нормальной» жизни; представления о «пороге бедности»; оценка изменения финансового состояния за последние три месяца;
- психосоциальные: удовлетворенность жизнью, самооценка уровня здоровья, уровень депрессивных состояний, взгляд на будущее/прошлое, трудовая мотивация (готовность много работать ради денег / меньше работать ради свободного времени), представления о важности собственных усилий;
 - интернет-активность (прокси цифровой грамотности): частота использования Интернета;
 - контекстные: продолжительность заполнения анкеты, готовность принимать участие в будущих волнах исследования.

Результаты

Доля несодержательных ответов на вопросы о доходах/расходах

Наибольшее затруднение вызвали вопросы о доходах и расходах семьи — на них затруднились ответить 16,5% и 13,6% респондентов соответственно (табл. 2). Всего 6,4% затруднились с ответом на вопрос об индивидуальном доходе, что значительно меньше, чем в подобных исследованиях за рубежом (14–35% в США [30], 20–40% в Европе [13]). Это может объясняться следующими особенностями дизайна опроса: настройкой маршрутизации вопросов о финансах только для работающих респондентов (сужает целевую аудиторию до людей с более стабильными и однозначно определяемыми доходами); наличием материального стимулирования участия, формирующего выборку с повышенной готовностью к сотрудничеству [9]; а также применением метода CAPI, при котором интервьюер снижает часть когнитивной нагрузки на респондента.

Таблица 2

Эффективность уточняющих вопросов

Измеряемый финансовый показатель	Количество опрошенных (N)	Доля несодержательных ответов на открытый вопрос, %	Доля содержательных ответов на уточняющий вопрос, от числа отвечавших на уточняющий вопрос, %	Доля несодержательных ответов после уточняющего вопроса, %
Индивидуальный доход	3792	6,4	36,1	4,1
Доходы семьи	6041	16,5	40,9	9,8
Расходы семьи	6041	13,6	33,7	9,1

Более трети изначально затруднившихся дали содержательный ответ на уточняющий вопрос. Лучше всего он сработал для измерения доходов семьи (40,9% из не указавших доходы в первом вопросе указали их во втором), менее успешно — для индивидуального дохода (36,1%) и семейных расходов (33,7%). Доля несодержательных ответов после уточняющего вопроса составляет 4,1–9,8% и близка к полученному результату в исследовании на выборке США (9,4% [30]).

***Характеристики респондентов,
повышающие/понижающие шансы выбора несодержательных ответов
на вопросы о доходах/расходах***

Результаты применения метода Random Forest с использованием алгоритма Boruta для выделения важных переменных позволили выделить основные переменные для включения в логистические регрессионные модели. В трех моделях с зависимыми переменными, отражающими выбор несодержательного ответа на открытый финансовый вопрос (личные доходы, доходы/расходы семьи), были выделены от 11 до 39 стабильно значимых предикторов, которые впоследствии включены в соответствующие логистические регрессии в качестве независимых переменных¹¹.

Анализ трех логистических регрессий показал (табл. 3), что склонность респондентов давать несодержательные ответы на финансовые вопросы систематически связана с рядом социально-демографических и поведенческих характеристик.

Во-первых, выделяются семейное положение и структура домохозяйства. Респонденты, состоящие в браке, но проживающие отдельно от партнера, демонстрируют наиболее высокую вероятность несодержательных ответов по всем моделям (рост шансов более чем в 2 раза по сравнению с женатыми, живущими вместе). Никогда не состоявшие в браке респонденты также чаще дают такие ответы, хотя эффект здесь умереннее и проявляется только в случае вопросов про доходы (личные и семейные). Напротив, разведенные демонстрируют меньшую склонность к несодержательным ответам в случае вопросов о личных доходах.

Во-вторых, важную роль играют занятость и трудовой статус. По сравнению с учащимися и одновременно работающими (базовая категория) большинство других групп — работающие, пенсионеры (работающие и неработающие), безработные, респонденты в отпуске по уходу за ребенком или занятые домашним хозяйством — реже дают несодержательные ответы. Наиболее выраженное снижение вероятности наблюдается среди работающих респондентов, в том числе пенсионеров. При этом в случае вопросов про семейные доходы и расходы вероятность несодержательного ответа выше у респондентов, которые учатся и не работают, по сравнению с теми, кто учится и работает.

В-третьих, выявлены значимые различия, связанные со здоровьем и психологическим состоянием. Наличие выраженных депрессивных состояний увеличивает риск несодержательных ответов на 18–32% в зависимости от модели.

¹¹ Таблицы с результатами всех моделей представлены в репозитории в файле «Результаты Random Forest Boruta (таблицы)» по ссылке: <https://disk.360.yandex.ru/d/PUeCaJDAtj2B2w>

Аналогично, плохое здоровье существенно повышает вероятность несодержательных ответов (от 50 до 86% по сравнению с респондентами с очень хорошим здоровьем).

В-четвертых, большое значение имеют характеристики финансового состояния. Наиболее сильным предиктором выступает осведомленность о расходах: респонденты, которые не знают расходов своей семьи, дают несодержательные ответы на вопросы о семейных доходах и расходах в 3–5 раз чаще, чем полностью осведомленные. Также повышают вероятность несодержательных ответов отсутствие сбережений, их сокращение за последние три месяца и самооценка уровня достатка как высокого. Более короткий горизонт финансового планирования («никаких планов» или «на несколько месяцев») также увеличивает вероятность несодержательных ответов.

Наконец, выявлены различия по территориальным и социально-демографическим характеристикам. Жители деревень и сел, малых и средних городов чаще дают несодержательные ответы, чем респонденты из крупнейших городов (Москва и Санкт-Петербург). Реже склонны уклоняться от ответа женщины и активные пользователи Интернета.

Таблица 3

Результаты логистических регрессий с зависимыми переменными, отражающими факт выбора несодержательного ответа на соответствующий финансовый вопрос

Предиктор	Личные доходы		Семейные доходы		Семейные расходы	
	OR	p-value	OR	p-value	OR	p-value
Количество членов семьи	0,838	0,040 *	1,093	0,045 *	0,963	0,456
Возраст респондента	1,022	0,005 **	1,013	0,003 **	1,006	0,210
Женат/замужем, живем порознь (база: женат/замужем, живем вместе)	3,826	0,034 *	2,887	0,013 *	2,650	0,030 *
Никогда не был(-а) в браке (база: женат/замужем, живем вместе)	0,463	0,003 **	1,387	0,010 *	1,037	0,795
Использование Интернета: почти каждый день (база: каждый день)	0,628	0,005 **	—	—	0,774	0,011 *
Учусь, не работаю (база: учусь и работаю)	—	—	1,763	0,012 *	2,069	0,004 **
Работаю, не на пенсии (база: учусь и работаю)	0,473	0,004 **	0,840	0,324	0,904	0,621
Работающий пенсионер (база: учусь и работаю)	0,293	0,001 ***	0,705	0,147	0,805	0,429
На пенсии, не работаю (база: учусь и работаю)	—	—	0,551	0,012 *	0,832	0,486
Депрессивные состояния (фактор)	1,311	0,000 ***	1,248	0,000 ***	1,182	0,000 ***
Здоровье: плохое (база: хорошее)	—	—	1,597	0,016 *	1,860	0,003 **
Осведомленность о расходах: скорее да (база: да)	0,749	0,253	1,481	0,004 **	1,615	0,001 **

Продолжение таблицы 3

Предиктор	Личные доходы		Семейные доходы		Семейные расходы	
	OR	p-value	OR	p-value	OR	p-value
Осведомленность о расходах: нет (база: да)	0,923	0,808	3,797	0,000 ***	4,864	0,000 ***
Изменение сбережений за 3 месяца: так же (база: больше)	1,649	0,127	1,645	0,002 **	1,581	0,010 *
Изменение сбережений за 3 месяца: меньше (база: больше)	1,896	0,109	2,019	0,000 ***	1,759	0,011 *
Изменение сбережений за 3 месяца: нет (база: больше)	3,021	0,003 **	2,467	0,000 ***	1,806	0,003 **
Горизонт планирования: никаких планов (база: 3 или больше лет)	1,977	0,001 ***	—	—	2,093	0,000 ***
Ожидания по материальному положению: так же (база: очень мало)	—	—	1,327	0,009 **	1,277	0,046 *
Количество машин: 1 (база: 0)	1,423	0,024 *	1,283	0,004 **	1,151	0,136
Размер населенного пункта: менее 100 тыс. (база: Москва и Санкт-Петербург)	1,179	0,515	1,266	0,094 ■	1,332	0,062 ■
Размер населенного пункта: деревня/село (база: Москва и Санкт-Петербург)	1,151	0,578	1,517	0,002 **	1,698	0,000 ***
Согласие на следующий этап	—	—	0,604	0,000 ***	0,690	0,000 ***

Примечание:

*** p-value < 0.001; ** p-value < 0.01; * p-value < 0.05; ■ p-value < 0.1.

Из-за большого количества независимых переменных приведены только те предикторы, которые являются значимыми хотя бы в одной из логистических регрессий. Символом «—» помечены переменные, которые не были включены в регрессию, так как не были определены методом Random Forest как потенциально важные предикторы.

OR — Odds Ratio (отношение шансов).

Ячейки со значениями отношений шансов, которые являются значимыми, выделены градиентной заливкой (от красного — значение, близкое к 0, к оранжевому — ближе к 1, к зеленому — значение сильно больше 1) для упрощения интерпретации.

***Характеристики респондентов,
повышающие/понижающие шансы выбора содержательных ответов
на уточняющие вопросы о доходах/расходах***

По результатам применения моделей Random Forest с алгоритмом Boruta для определения важности предикторов содержательных ответов на уточняющие вопросы относительно финансов (личные доходы, доходы/расходы семьи) были выделены ключевые предикторы, стабильно показавшие связь с зависимыми пе-

ременными во всех итерациях построения моделей случайного леса¹². По итогам были выделены от 10 до 21 независимой переменной для последующего использования в моделях логистической регрессии в качестве независимых переменных. Результаты логистических регрессий, описывающих вероятность выбора содержательного ответа на уточняющие вопросы относительно финансов (личные доходы, доходы/расходы семьи), представлены в таблице 4.

Во-первых, наблюдается связь между выбором содержательного ответа на уточняющий вопрос с социально-демографическими характеристиками. Респонденты из более многочисленных семей и с более высоким уровнем образования реже предоставляют содержательный ответ на уточняющий вопрос. Аналогичный эффект наблюдается у респондентов с разными категориями занятости. Кроме того, жители городов среднего и малого размера чаще дают содержательные ответы относительно доходов семьи, чем жители крупных городов.

Во-вторых, финансовая ситуация и поведение также связаны с вероятностью выбора содержательного ответа на уточняющий вопрос. Отсутствие сбережений повышает готовность такого выбора. Напротив, финансовая тревожность снижает вероятность успешного уточнения в вопросах как о доходах, так и о расходах. Осведомленность о расходах играет одну из ключевых ролей в эффективности уточняющих вопросов о расходах семьи: неполная осведомленность (ответ «скорее да») резко снижает шансы дать содержательный ответ на уточняющий вопрос по сравнению с полной осведомленностью.

В-третьих, наблюдается также эффект времени заполнения анкеты: его увеличение связано с меньшей вероятностью предоставить содержательный ответ на уточняющий вопрос о личных доходах.

Таблица 4

Результаты логистических регрессий с зависимыми переменными, отражающими факт выбора содержательного ответа на соответствующий уточняющий финансовый вопрос после выбора несодержательного ответа на открытый вопрос

Предиктор	Личные доходы		Семейные доходы		Семейные расходы	
	OR	p-value	OR	p-value	OR	p-value
Количество членов семьи	0,81	0,078 ■	1,08	0,335	—	—
Полное среднее образование (база: начальное/неполное среднее образование)	0,00	0,992	0,51	0,079 ■	—	—
Начальное проф. образование (ПТУ) (база: начальное/неполное среднее образование)	0,00	0,992	0,39	0,040 *	—	—
Среднее проф. образование (техникум) (база: начальное/неполное среднее образование)	0,00	0,992	0,55	0,097 ■	—	—

¹² Таблицы с результатами всех моделей представлены в репозитории в файле «Результаты Random Forest Boruta (таблицы)» по ссылке: <https://disk.360.yandex.ru/d/PUeCaJDaTj2B2w>

Продолжение таблицы 4

Предиктор	Личные доходы		Семейные доходы		Семейные расходы	
	OR	p-value	OR	p-value	OR	p-value
Незаконченное высшее (база: начальное/неполное среднее образование)	0,00	0,992	0,29	0,017 *	—	—
Высшее образование (база: начальное/неполное среднее образование)	0,00	0,992	0,44	0,027 *	—	—
Учусь и не работаю (база: учусь и работаю)	—	—	2,17	0,059 ■	—	—
Осведомлен о расходах: скорее да (база: да)	—	—	0,73	0,175	0,6	0,007 **
Доля расходов на еду	—	—	1,01	0,095 ■	1,01	0,11
Изменение сбережений: нет сбережений (база: больше)	—	—	2,87	0,007 **	1,34	0,465
Финансовая тревожность	—	—	0,93	0,002 **	0,92	0,001 ***
Удовлетворенность жильем: скорее да (база: да)	—	—	0,57	<0,001 ***	—	—
Размер НП: 500–999 тыс. (база: Москва и Санкт-Петербург)	—	—	3,64	<0,001 ***	—	—
Размер НП: 250–499 тыс. (база: Москва и Санкт-Петербург)	—	—	1,81	0,073 ■	—	—
Размер НП: менее 100 тыс. (база: Москва и Санкт-Петербург)	—	—	1,58	0,092 ■	—	—
Время заполнения анкеты (в минутах)	0,99	0,096 ■	1	0,1	—	—

Примечание:

*** p-value < 0.001; ** p-value < 0.01; * p-value < 0.05; ■ p-value < 0.1

Из-за большого количества независимых переменных приведены только те предикторы, которые являются значимыми хотя бы в одной из логистических регрессий. Символом «—» помечены переменные, которые не были включены в регрессию, так как не были определены методом Random Forest как потенциально важные предикторы.

OR — Odds Ratio (отношение шансов).

Ячейки со значениями отношений шансов, которые являются значимыми, выделены градиентной заливкой (от красного — значение, близкое к 0, к оранжевому — ближе к 1, к зеленому — значение сильно больше 1) для упрощения интерпретации.

Различия в наборах предикторов выбора несодержательных ответов на вопросы о доходах/расходах и содержательных ответов на уточняющие вопросы

Наборы значимых предикторов несодержательных ответов на первоначальные и уточняющие вопросы частично совпадают, но не являются полностью идентичными. Совпадения есть прежде всего для базовых социально-демографических характеристик и некоторых установочных и поведенческих переменных, связанных с финансовой грамотностью и уверенностью в собственных оценках. Различия в наборах обусловлены несколькими причинами.

Зависимыми переменными в моделях являются разные по смыслу переменные. В первом случае (табл. 3) описывается модель с предсказанием несодержательного ответа на открытый вопрос о финансах, во втором (табл. 4) — представлены результаты модели, предсказывающей предоставление валидного ответа на закрытый уточняющий вопрос.

Первоначальные и уточняющие вопросы различаются по когнитивной нагрузке. В первом случае вопрос сложнее, поскольку требует более точного воспоминания (нужно указать конкретную сумму в рублях, а не выбрать из представленного набора вариантов, как в случае уточняющего вопроса).

Моделирование проводится на разных выборках. Уточняющие вопросы задаются подвыборке респондентов — тем, кто выбрал вариант «Затрудняюсь ответить» в открытом вопросе. Следовательно, анализ для уточняющих вопросов проводится на более однородной группе, где вариативность некоторых предикторов снижается, а значимость других, наоборот, возрастает.

Ограничения исследования

Настоящее исследование имеет ряд методологических ограничений, которые необходимо учитывать при интерпретации результатов.

Во-первых, в данной работе не проводилось разделение несодержательных ответов на категории «Затрудняюсь ответить» и «Отказ от ответа», поскольку в анкете они были объединены в одну опцию. Между тем исследования показывают, что эти типы ответов обусловлены разными механизмами: «Затрудняюсь ответить» обычно связано с недостатком знаний или когнитивными трудностями, тогда как «Отказ от ответа» чаще отражает восприятие вопроса как слишком чувствительного [7; 25]. Различение этих ответов позволило бы более точно определить, для каких категорий респондентов и в связи с каким механизмом действия уточняющие вопросы более или менее эффективны.

Во-вторых, когнитивные механизмы выбора несодержательных ответов рассматривались на основе вторичных поведенческих индикаторов, а не прямых измерений когнитивной нагрузки. Для более определенных выводов потребуются дополнительные измерения воспринимаемой сложности вопросов, например с использованием шкалы Пааса для оценки когнитивной нагрузки [23] и подходов когнитивного интервью [2], а также отдельная оценка готовности респондентов открыто говорить о финансах.

В-третьих, настоящее исследование не оценивает качество данных, полученных с помощью уточняющих вопросов. Остается открытым вопрос, насколько точны ответы респондентов в интервальном формате по сравнению с точными числовыми значениями. Для решения этой задачи потребовались бы валидационные исследования с использованием административных данных или замеров с помощью других методик¹³ [27].

Наконец, используемая методология Random Forest с алгоритмом Boruta, хотя и позволяет эффективно отбирать значимые предикторы [10; 19], имеет ограничения в интерпретации причинно-следственных связей. Выявленные предикторы следует рассматривать как связи, а не как каузальные факторы.

Дискуссия

Снижение отклика на опросы и рост затрат на сбор данных, в том числе удорожание процедур дозвона, стимулируют исследователей искать более экономичные способы удержания респондентов. Особенно это актуально для опросов, затрагивающих чувствительные темы, в частности, вопросы о доходах и расходах.

Несодержательные ответы на вопросы о финансах снижают общее качество опросных данных: ведут к искажению статистики доходов и расходов¹⁴, а также к потере или искажению информации об изучаемых характеристиках респондентов, для которых такие вопросы оказываются слишком личными или когнитивно сложными [14].

Проведенное исследование проверяет и подтверждает эффективность методики уточняющих (*probing*) вопросов как инструмента снижения доли неответов на финансовые вопросы в российских опросах. До 41% респондентов, изначально отказавшихся от ответа, предоставили содержательную информацию после уточняющего вопроса с заданными интервалами. Этот результат ниже аналогичных показателей, полученных в США (до 56%) [30], однако демонстрирует устойчивую эффективность подхода.

Наибольшая результативность уточняющих вопросов наблюдается в отношении семейных доходов (41%), тогда как для индивидуальных доходов она ниже (36%). Это указывает на то, что основным механизмом их действия является снижение когнитивной нагрузки, а не смягчение чувствительности. Вопросы о семейных финансах требуют агрегирования информации от нескольких источников, более сложных вычислений и информированности, что создает дополнительные когнитивные барьеры [11; 13].

Анализ социально-демографических характеристик, влияющих на склонность к неответам, подтверждает преобладание фактора когнитивной сложности. Наиболее сильные предикторы несодержательных ответов связаны с ограниченными когнитивными ресурсами и доступностью информации: проживание отдельно

¹³ Using administrative data to assess nonresponse bias in the Current Population Survey 2023. Annual Social and Economic Supplement. U.S. Census Bureau. — URL: <https://www.census.gov/newsroom/blogs/research-matters/2023/09/using-administrative-data-nonresponse-cps-asec.html> (дата обращения 20.10.2025).

¹⁴ Ibid.

от супруга (неосведомленность о его доходах), плохое здоровье, депрессивные состояния, короткий горизонт финансового планирования и отсутствие сбережений. Реже уклоняются от ответов работающие респонденты (особенно работающие пенсионеры), а также те, кто согласился участвовать в следующей волне исследования, что указывает на их более высокую мотивацию к сотрудничеству.

Тем не менее результаты подтверждают, что вопросы о финансах являются преимущественно сенситивными для респондентов. Об этом свидетельствует прежде всего ограниченная эффективность уточняющих вопросов: более половины из тех, кто уклонился от первоначального ответа, не предоставили данные. Кроме того, одним из ключевых предикторов содержательных ответов на уточняющие вопросы является уровень образования, который, как правило, повышает шансы на содержательный ответ [13; 14; 30]. Однако в нашем исследовании более образованные респонденты, напротив, реже указывали финансовую информацию в ответ на уточняющий вопрос, что может быть связано с осознанным контролем над раскрываемой информацией.

Место проживания, часто имеющее значение при формировании выборки, также связано с уклонением от ответов на финансовые вопросы. Жители малых населенных пунктов чаще избегают ответов, тогда как респонденты из средних городов (500–999 тыс. человек) в 3,6 раза чаще, чем москвичи и петербуржцы, предоставляют данные в ответ на уточняющие вопросы. Вероятно, мотивация жителей средних городов предоставлять финансовые данные находится между ограничениями, характерными для малых населенных пунктов (низкая цифровая грамотность, неформализованность доходов), и культурой приватности, присущей крупнейшим городам.

Полученные результаты позволяют сформулировать несколько практических рекомендаций для улучшения дизайна опросов, включающих финансовые вопросы.

— *Использование уточняющих вопросов* — недорогой и эффективный способ сокращения доли неответов до 4–10%. Это особенно важно для исследований с квотной или стратифицированной выборкой, где компенсация потери респондентов может требовать значительного повышения бюджета исследования, а также для исследований, задачи которых состоят в получении косвенных оценок доходов и расходов с помощью опросных данных в дополнение к статистическим, которые могут быть искажены из-за теневой экономики и других особенностей.

— *Адаптация формулировок* для различных групп респондентов, оказавшихся в группе риска неответов. Например, для работающих и более образованных целесообразно запрашивать сведения по отдельным источникам дохода с последующим суммированием исследователями.

— *Использование фильтров* в вопросах о семейных финансах, особенно для респондентов, состоящих в неформальных браках или живущих порознь.

— *Усиление гарантий конфиденциальности* в преамбуле блока вопросов, что может способствовать снижению опасения относительно разглашения персональных данных.

— *Статистическая корректировка смещений* с учетом факторов, влияющих на уклонение от ответов, например с использованием оценки Хорвича — Томпсона [14].

СВЕДЕНИЯ ОБ АВТОРАХ

Шмелева Евгения Дмитриевна — кандидат наук об образовании, заместитель директора, Центр внутреннего мониторинга, НИУ ВШЭ. **Телефон:** +7 (925) 365-51-53. **Электронная почта:** eshmeleva@hse.ru SPIN-код: 9935-5193

Лебедев Даниил Вадимович — кандидат социологических наук, аналитик, Центр внутреннего мониторинга, НИУ ВШЭ. **Телефон:** +7 (909) 155-60-28. **Электронная почта:** dlebedev@hse.ru SPIN-код: 1659-9524

Брагина Екатерина Викторовна — аналитик, Центр внутреннего мониторинга, НИУ ВШЭ. **Телефон:** +7 (921) 828-93-99. **Электронная почта:** e.bragina@hse.ru

SOTSIOLOGICHESKIY ZHURNAL = SOCIOLOGICAL JOURNAL. 2026. Vol. 32. No. 3. P. 9–29. DOI: 10.19181/socjour.2026.32.3.1

Research Article

EVGENIYA D. SHMELEVA¹, DANIIL V. LEBEDEV¹, EKATERINA V. BRAGINA¹

¹ HSE University.

11 Pokrovsky Boulevard, 109028, Moscow, Russian Federation

**“BUT PERHAPS...”: THE EFFECTIVENESS OF PROBING QUESTIONS
FOR REDUCING ITEM NON-RESPONSE TO FINANCIAL QUESTIONS**

Abstract. A high proportion of non-substantive responses to questions about income and expenditures compromises data quality, limits the generalizability of findings, and can lead to sampling bias. The objective of this study is to assess the effectiveness of using follow-up questions to improve the quality of responses to financial questions. The study is based on data from a representative survey of the adult population of Russia, conducted by HSE University as part of the project “Household Economic Behavior Monitoring” (Spring 2025, N = 6 041). The first stage involved identifying predictors of non-substantive responses to the financial questions, while the second stage assessed the effectiveness of follow-up (probing) questions for different respondent categories. Predictors were selected using the Random Forest method with the Boruta algorithm and logistic regressions. The results show that the share of non-substantive responses was 6% for individual income, 17% for household income, and 14% for household expenditures. Follow-up questions converted 36–41% of the initial non-responses into substantive answers, thereby reducing the overall non-response rate to 4–10%. The highest risk of non-response was identified among respondents living separately from their spouse (a more than two-fold increase in the odds), those unaware of household expenditures (a 3- to 5-fold increase), those with depressive states (18–32% increase), poor health (50–86% increase), no savings, a short planning horizon, and residents of small towns. Follow-up questions were most effective for respondents with lower education levels, those unaware of expenditures, those without savings, and those from small towns. The method is only partially effective: while it reduces cognitive complexity it does not affect the sensitivity of income-related questions. Based on these findings, recommendations are proposed for improving the design of surveys that include financial questions.

Keywords: survey data quality; item nonresponse; do not know answers; sensitive questions; probing questions; financial questions; income; expenditures.

For citation: Shmeleva, E.D., Lebedev, D.V., Bragina, E.V. “But perhaps...”: The Effectiveness of Probing Questions for Reducing Item Non-response to Financial Questions. *Sotsiologicheskii Zhurnal = Sociological Journal*. 2026. Vol. 32. No. 3. P. 9–29. DOI: 10.19181/socjour.2026.32.3.1

Acknowledgements. This work is an output of project HSE-MS-2025-006 implemented at HSE University.

INFORMATION ABOUT THE AUTHORS

Evgeniya D. Shmeleva — Candidate of Pedagogical Sciences, Deputy Director, Centre for Institutional Research, HSE University. **Phone:** +7 (925) 365-51-53. **Email:** eshmeleva@hse.ru ORCID: 0000-0001-8004-3315 ResearcherID: P-9230-2019

Daniil V. Lebedev — Candidate of Sociological Sciences, Analyst, Centre for Institutional Research, HSE University. **Phone:** +7 (909) 155-60-28. **Email:** dlebedev@hse.ru ORCID: 0000-0001-7246-8022 ResearcherID: D-6656-2019

Ekaterina V. Bragina — Analyst, Centre for Institutional Research, HSE University. **Phone:** +7 (921) 828-93-99. **Email:** e.bragina@hse.ru

ЛИТЕРАТУРА / REFERENCES

1. *Аласшеев С.Ю.* Недостоверность ответов респондентов на вопросы о доходах // Социологический журнал. 2015. Том 21. № 3. С. 29–44. DOI: [10.19181/socjour.2015.21.3.2376](https://doi.org/10.19181/socjour.2015.21.3.2376) EDN: UKAKEZ
Alasheev S.Yu. Inaccuracy of respondents' answers to income questions. *Sotsiologicheskii Zhurnal = Sociological Journal*. 2015. Vol. 21. No. 3. P. 29–44. DOI: [10.19181/socjour.2015.21.3.2376](https://doi.org/10.19181/socjour.2015.21.3.2376) (In Russ.)
2. *Девятко И.Ф., Лебедев Д.В.* Глазами интервьюера, глазами респондента: контуры нового подхода к оценке когнитивной нагрузки при проведении опроса // Мониторинг общественного мнения: экономические и социальные перемены. 2017. № 5. С. 1–19. DOI: [10.14515/monitoring.2017.5.01](https://doi.org/10.14515/monitoring.2017.5.01) EDN: YLYNWJ
Deviatko I.F., Lebedev D.V. Through the eyes of interviewer and respondent: towards a new approach to measuring cognitive load in surveys. *Monitoring obshchestvennogo mneniya: Ekonomicheskie i sotsial'nye peremeny*. 2017. No. 5. P. 1–19. DOI: [10.14515/monitoring.2017.5.01](https://doi.org/10.14515/monitoring.2017.5.01) (In Russ.)
3. *Мягков А.Ю.* Косвенные методы в сенситивных исследованиях. М.: ИНФРА-М, 2024. — 245 с. DOI: [10.12737/2110854](https://doi.org/10.12737/2110854) EDN: KWXANI
Myagkov A.Yu. *Indirect methods in sensitive research*. Moscow: INFRA-M publ., 2024. 245 p. DOI: [10.12737/2110854](https://doi.org/10.12737/2110854) (In Russ.)
4. *Мягков А.Ю.* Сенситивные исследования: опыт ретроспективного анализа и концептуализации // Социологический журнал. 2023. Том 29. № 3. С. 8–28. DOI: [10.19181/socjour.2023.29.3.1](https://doi.org/10.19181/socjour.2023.29.3.1) EDN: BKJOVE
Myagkov A.Yu. Sensitive research: retrospective analysis and conceptualization *Sotsiologicheskii Zhurnal = Sociological Journal*. 2023. Vol. 29. No. 3. P. 8–28. DOI: [10.19181/socjour.2023.29.3.1](https://doi.org/10.19181/socjour.2023.29.3.1) (In Russ.)
5. *Юдин Г.Б.* Территориальная локализация и уровень неответов в массовом опросе // Социологический журнал. 2008. № 1. С. 49–72. Дата обращения 20.10.2025. URL: <https://www.journal-socjournal.ru/index.php/socjour/article/view/983/937>¹⁵
Yudin G.B. Territorial localization and nonresponse rate in mass surveys. *Sotsiologicheskii Zhurnal = Sociological Journal*. 2008. No. 1. P. 49–72. Accessed 20.10.2025. URL: <https://www.journal-socjournal.ru/index.php/socjour/article/view/983/937> (In Russ.)
6. Breiman L. Random forests. *Machine Learning*. 2001. Vol. 45. No. 1. P. 5–32. DOI: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324)
7. Converse J.M. Predicting No Opinion in the Polls. *Public Opinion Quarterly*. 1976. Vol. 40. No. 4. P. 515–530. DOI: [10.1086/268337](https://doi.org/10.1086/268337)
8. Davern M., Rodin H., Beebe T.J., Call K.T. The Effect of Income Question Design in Health Surveys on Family Income, Poverty and Eligibility Estimates. *Health Services Research*. 2005. Vol. 40. No. 5. Pt. 1. P. 1534–1552. DOI: [10.1111/j.1475-6773.2005.00416.x](https://doi.org/10.1111/j.1475-6773.2005.00416.x)

¹⁵ Автор статьи признан иностранным агентом.

9. De Leeuw E.D. Reducing Missing Data in Surveys: An Overview of Methods. *Quality & Quantity*. 2001. Vol. 35. P. 147–160. DOI: [10.1023/A:1010395805406](https://doi.org/10.1023/A:1010395805406)
10. Degenhardt F., Seifert S., Szymczak S. Evaluation of variable selection methods for random forests and omics data sets. *Briefings in Bioinformatics*. 2019. Vol. 20. No. 2. P. 492–503. DOI: [10.1093/bib/bbx124](https://doi.org/10.1093/bib/bbx124)
11. Hansen K., Kneale D. Does How You Measure Income Make a Difference to Measuring Poverty? Evidence from the UK. *Social Indicators Research*. 2013. Vol. 110. P. 1119–1140. DOI: [10.1007/s11205-011-9976-5](https://doi.org/10.1007/s11205-011-9976-5)
12. Heeringa S.G., Hill D.H., Howell D.A. *Unfolding Brackets for Reducing Item Nonresponse in Economic Surveys. AHEAD/HRS Report No. 94–029*. Ann Arbor: Survey Research Center, Institute for Social Research, University of Michigan, 1993. Accessed 20.10.2025. URL: <https://hrs.isr.umich.edu/publications/biblio/10985>
13. Jabkowski P., Piekut A. Between Task Complexity and Question Sensitivity: Nonresponse to the Income Question in the 2008–2018 European Social Survey. *Survey Research Methods*. 2024. Vol. 18. No. 2. P. 113–135. DOI: [10.18148/srm/2024.v18i2.8117](https://doi.org/10.18148/srm/2024.v18i2.8117)
14. Jabkowski P., Piekut A. Not Random and Not Ignorable. An Examination of Nonresponse to Income Question in the European Social Survey, 2008–2018. *Public Opinion Quarterly*. 2023. Vol. 36. No. 3. DOI: [10.1177/1525822X231194178](https://doi.org/10.1177/1525822X231194178)
15. Jansen W., Verhoeven W.J., Dessens J., Róbert P. The long and short of asking questions about income: a comparison using data from Hungary. *Quality & Quantity*. 2013. Vol. 47. No. 4. P. 1957–1969. DOI: [10.1007/s11135-011-9636-5](https://doi.org/10.1007/s11135-011-9636-5)
16. Kaminska O., McCutcheon A.L., Billiet J. Satisficing Among Reluctant Respondents in a Cross-National Context. *Public Opinion Quarterly*. 2010. Vol. 74. No. 5. P. 956–984. DOI: [10.1093/poq/nfq062](https://doi.org/10.1093/poq/nfq062)
17. Krosnick J.A. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*. 1991. Vol. 5. No. 3. P. 319–335. DOI: [10.1002/acp.2350050305](https://doi.org/10.1002/acp.2350050305)
18. Kuha J., Butt S., Katsikatsou M., Skinner C.J. The Effect of Probing “Don’t Know” Responses on Measurement Quality and Nonresponse in Surveys. *Journal of Official Statistics*. 2018. Vol. 34. No. 1. P. 26–40. DOI: [10.1080/01621459.2017.1323640](https://doi.org/10.1080/01621459.2017.1323640)
19. Kurs M.B., Rudnicki W.R. Feature Selection with the Boruta Package. *Journal of Statistical Software*. 2010. Vol. 36. No. 11. P. 1–13. DOI: [10.18637/jss.v036.i11](https://doi.org/10.18637/jss.v036.i11)
20. Lepkowski J. Non-observation error in household surveys in developing countries. *Household Surveys in Developing and Transition Countries*. Ed. by Department of Economic and Social Affairs, Statistics Division. N.Y.: United Nations, 2005. P. 149–169. Accessed 20.10.2025. URL: https://unstats.un.org/unsd/hhsurveys/pdf/Chapter_8.pdf
21. Lusardi A., Mitchell O.S. The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature*. 2014. Vol. 52. No. 1. P. 5–44. DOI: [10.1257/jel.52.1.5](https://doi.org/10.1257/jel.52.1.5)
22. Mignogna G., Carey C.E., Wedow R. et al. Patterns of item nonresponse behaviour to survey questionnaires are systematic and associated with genetic loci. *Nature Human Behaviour*. 2023. Vol. 7. P. 1371–1387. DOI: [10.1038/s41562-023-01632-7](https://doi.org/10.1038/s41562-023-01632-7)
23. Paas F.G. Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*. 1992. Vol. 84. No. 4. P. 429–434. DOI: [10.1037/0022-0663.84.4.429](https://doi.org/10.1037/0022-0663.84.4.429)
24. Pearce-Morris J., Choi S.-W., Roth V., Young R. Substantive Meanings of Missing Data in Family Research: Does “Don’t Know” Matter? *Journal of Family Issues*. 2014. Vol. 35. No. 5. P. 665–690. DOI: [10.1080/01494929.2014.938292](https://doi.org/10.1080/01494929.2014.938292)

25. Riphahn R.T., Serfling O. Item non-response on income and wealth questions. *Empirical Economics*. 2005. Vol. 30. No. 2. P. 521–538. DOI: [10.1007/s00181-005-0247-7](https://doi.org/10.1007/s00181-005-0247-7)
26. Rothacher Y., Strobl C. Identifying Informative Predictor Variables With Random Forests. *Journal of Educational and Behavioral Statistics*. 2024. Vol. 49. No. 4. P. 595–629. DOI: [10.3102/10769986231193327](https://doi.org/10.3102/10769986231193327)
27. Schwartz L., Paulin G. Improving response rates to income questions. *American Statistical Association (ASA) Section on Survey Research Methods Proceedings*. 2000. P. 965–970. Accessed 20.10.2025. URL: http://www.asasrms.org/Proceedings/papers/2000_165.pdf
28. Tourangeau R., Yan T. Sensitive Questions in Surveys. *Psychological Bulletin*. 2007. Vol. 133. No. 5. P. 859–883. DOI: [10.1037/0033-2909.133.5.859](https://doi.org/10.1037/0033-2909.133.5.859)
29. Wang K. Effects of unfolding brackets on the quality of income data in a telephone survey. *Survey Practice*. 2010. Vol. 3. No. 3. P. 1–7. DOI: [10.29115/SP-2010-0014](https://doi.org/10.29115/SP-2010-0014)
30. Yan T., Curtin R., Jans M. *The Impact of Unit Nonresponse on Income Nonresponse*. Ann Arbor: Survey Research Center, University of Michigan, 2006. 22 p. Accessed 20.10.2025. URL: <https://data.sca.isr.umich.edu/fetchdoc.php?docid=32879>

Статья поступила в редакцию: 20.10.2025; поступила после рецензирования и доработки: 16.04.2026; принята к публикации: 11.05.2026.

Received: 20.10.2025; revised after review: 16.04.2026; accepted for publication: 11.05.2026.