

Прикладная эконометрика, 2022, т. 66, с. 118–143.

Applied Econometrics, 2022, v. 66, pp. 118–143.

DOI: 10.22394/1993-7601-2022-66-118-143

Ю. В. Раскина, О. А. Подкорытова, Р. К. Кучаков¹

Детерминанты здоровья и гетерогенность ответов при самооценке здоровья в России: подход якорных виньеток

С использованием метода якорных виньеток и модели иерархического упорядоченного выбора в статье показано, что во всех 7 аспектах здоровья, представленных в данных российской выборки 1-й волны SAGE, наблюдается гетерогенность самооценок здоровья. Учет этой гетерогенности существенно меняет значимость потенциальных детерминант здоровья, таких как образование, перманентный доход, место проживания и семейный статус.

Ключевые слова: детерминанты здоровья; гетерогенность ответов при самооценке здоровья; якорные виньетки; HOPIT; Россия; SAGE.

JEL classification: C21; I12.

1. Введение

1.1. Детерминанты здоровья

Здоровье людей определяют не только ресурсы медицинской системы. Эффект от медицинской помощи может быть усилен или ослаблен воздействием различных факторов (Marmot et al., 2008). В настоящей статье исследуется взаимосвязь между здоровьем индивида и аспектами его социально-экономического статуса. Экономисты обычно выделяют среди этих аспектов образование и экономическое благосостояние. Как правило, более образованные люди имеют более высокие показатели здоровья и меньший риск смертности. Это связано не только с лучшими условиями труда и медицинского страхования. Образованные люди предпринимают больше усилий по сохранению здоровья, могут эффективнее искать и использовать медицинскую информацию, добиваться получения качественной медицинской помощи, понимать и придерживаться схем лечения, предложенных врачом, а их социальное окружение, в свою очередь, может препятствовать формированию разрушительных и поддерживать благоприятные для здоровья привычки (Cutler,

¹ Раскина Юлия Владимировна — Европейский университет в Санкт-Петербурге, Санкт-Петербург; raskina@eu.spb.ru.

Подкорытова Ольга Анатольевна — Европейский университет в Санкт-Петербурге, Санкт-Петербургский государственный университет, Санкт-Петербург; opodkorytova@eu.spb.ru.

Кучаков Руслан Камилович — Институт проблем правоприменения при Европейском университете в Санкт-Петербурге, Санкт-Петербург; rkuchakov@eu.spb.ru.

Lleras-Muney, 2006). Важны также и другие характеристики: возраст и пол (Martinson, 2012; Stafford et al., 2005), расовая или этническая принадлежность (Krieger et al., 2003), регион проживания (Bambra, Popham, 2010; Montez, Berkman, 2013), проживание в сельской или городской местности (Miao, Wu, 2016; Riva et al., 2009).

Обычным инструментом для нахождения детерминант здоровья является оценка регрессионной модели, в которой зависимая переменная — мера здоровья, а независимые переменные — компоненты социально-экономического статуса и иные переменные, которые могут определять форму искомой зависимости. Для проведения подобных исследований необходима хорошая мера (или несколько мер) здоровья.

1.2. Самооценки здоровья

Самооценка здоровья является, пожалуй, наиболее часто используемой мерой здоровья в исследованиях, использующих данные для домохозяйств и индивидов. Эта мера охватывает всю телесную информацию, которая непосредственно доступна только самому человеку (Jylhä, 2009). Самооценка здоровья учитывает даже те его аспекты, которые остаются неохваченными при использовании каких-либо отдельных объективных показателей здоровья или биомаркеров (Герри, Байдин, 2017). Самооценка здоровья имеет статистически значимую корреляцию с объективными показателями здоровья (Jylhä et al., 2006; Manor et al., 2001), она является статистически значимым предиктором как использования медицинских услуг (Doorslaer, Koolman, 2004), так и смертности (DeSalvo et al., 2006; Idler, Benyamini, 1997; Nielsen et al., 2008).

Другая привлекательная черта самооценки здоровья заключается в том, что в массовых обследованиях населения ее получение гораздо дешевле, чем сбор комплексной информации о здоровье населения, проведение инструментальных и лабораторных исследований.

Расширением общего вопроса «Как вы оцениваете свое здоровье в целом?» является класс самооценок здоровья, учитывающих его отдельные аспекты. Всемирная организация здравоохранения разработала для Всемирного обследования состояния здоровья (World Health Survey) опросник по восьми аспектам (domains) здоровья: мобильность, забота о себе, боль и скованность в теле, когнитивные способности, межличностные взаимодействия, зрение, сон и энергия, эмоциональное состояние².

1.3. Гетерогенность ответов при самооценке здоровья

Казалось бы, есть прекрасный инструмент для измерения здоровья — его самооценка. Однако ее использование как меры истинного здоровья несет в себе потенциальную проблему. Многочисленные исследования показывают, что различия в нормах и ожиданиях людей в отношении собственного здоровья могут приводить к систематическому смещению ответов при самооценке здоровья у различных групп населения. Представители разных групп могут иметь разную интерпретацию того, что означает «плохое» или «отличное»

² Каждый аспект содержит два индикатора, которым соответствует два вопроса. Подробнее см. в разделе 3 «Данные и переменные».

здоровье, сильные или умеренные проблемы в том или ином аспекте здоровья. Говоря чуть формальнее, группы могут отличаться наборами пороговых точек между соседними категориями шкалы ответов (например, хорошее / очень хорошее здоровье) для оценки ненаблюдаемого истинного здоровья.

В англоязычной литературе для обозначения подобного смещения используются различные термины: state-dependent reporting errors (Kerkhofs, Lindeboom, 1995), scale of reference bias (Groot, Den Brink, 2007), differential item functioning (King et al., 2004; Scott et al., 2010), reporting heterogeneity bias (Bago d'Uva et al., 2008; Lindeboom, Doorslaer, 2004). В русскоязычной литературе предложен термин «гетерогенность ответов при самооценке здоровья» (Кучаков, Раскина, 2017; Герри, Байдин, 2017; Канева, Байдин, 2018), этот термин и будет использоваться в данной статье. Подчеркнем, что гетерогенность ответов при самооценке становится проблемой тогда, когда она не случайна, а если представители разных групп населения систематически формируют свое мнение о состоянии здоровья по-разному.

Предыдущие исследования свидетельствуют, что гетерогенность ответов при самооценке здоровья систематически наблюдается между странами, индивидами с разными доходами, уровнями образования, пола, возраста, проживающими в городе и в сельской местности, см., например, (Bago d'Uva et al., 2008; Jürges, 2007; Lindeboom, Doorslaer, 2004; Oksuzyan et al., 2019; Schneider et al., 2012).

К сожалению, недостаточно провести исследование для одной страны, т. к. специфика страны может сказаться на том, какие именно факторы влияют на здоровье. Например, в исследовании (Eikemo et al., 2008) было найдено, что в скандинавских странах, славящихся своей универсальной и щедрой социальной политикой, неравенство в отношении здоровья, обусловленное доходом, ниже, чем в англо-саксонских странах. В работе (Doorslaer, Koolman, 2004) найдена существенная разница в величине эффекта дохода на здоровье в 13 странах Евросоюза.

В статье (Lokshin, Ravallion, 2008) на данных RLMS-HSE³ было найдено, что самооценка здоровья взрослых россиян слабо связана с потреблением домохозяйства или индивидуальным доходом. Однако, если использовать более объективные показатели здоровья (испытываемые проблемы при физической активности и наличие болезненных симптомов), то эффект от дохода становится гораздо более выраженным. При этом он становится незначимым при контроле на возраст, образование и место жительства. Горшков и Жаринова (Gorshkov, Zharinova, 2016) на данных SAGE⁴ и RLMS-HSE показали, что более образованные люди имеют лучшее здоровье, выраженное его самооценкой, а также ведут более здоровый образ жизни; этот эффект сохраняется и после контроля на демографические переменные и доход. Исследование (Paul, Valtonen, 2016)⁵, использующее данные RLMS-HSE,

³ «Российский мониторинг экономического положения и здоровья населения НИУ ВШЭ (RLMS HSE)», проводимый Национальным исследовательским университетом «Высшая школа экономики» и ООО «Демоскоп» при участии Центра народонаселения Университета Северной Каролины в Чапел Хилле и Института социологии Федерального научно-исследовательского социологического центра РАН. Сайты обследования RLMS HSE: <https://rlms-hse.cpc.unc.edu> и <http://www.hse.ru/rlms>.

⁴ WHO Study on global AGEing and adult health, www.who.int/healthinfo/sage.

⁵ Работа (Пол и др., 2019), имеющая расширенный состав авторов и выборку, в целом повторяет данное исследование и его выводы.

показало, что различия в здоровье между более обеспеченными и менее обеспеченными индивидами уменьшались в период 1994–2012 гг.

Результаты (Канева, Байдин, 2018) свидетельствуют, что гетерогенность ответов при самооценке существует и зависит от пола, семейного положения, образования, места проживания и возраста. Недавнее исследование (Хабибуллина и др., 2021) также дает основания предполагать наличие такой гетерогенности. Исследователи находят, что при использовании самооценок здоровья в качестве меры истинного здоровья связь между здоровьем и образованием существует, но она пропадает при использовании более объективных показателей здоровья.

Как правило, более уязвимые группы населения систематически оценивают свое здоровье выше (демонстрируют оптимизм в оценке своего здоровья), чем менее уязвимые, при одном и том же уровне истинного здоровья — бедные выше оценивают свое здоровье, чем богатые, менее образованные — выше, чем более образованные. Пожилые часто оптимистичнее молодых в оценках своего здоровья. Объяснение этого исследователи видят в том, что люди сравнивают себя с теми, кто их окружает, со своей референтной группой (Molina, 2017). Так как бедные как правило менее здоровы, чем богатые, менее образованные — чем более образованные, а пожилые — чем молодые, то более низкий уровень их здоровья может чаще рассматриваться ими как нормальный или хороший. Получается, что измеренное различие в самооценке здоровья представляет собой комбинацию состояния здоровья с характеристиками индивида и смещением самооценки, связанной с этими характеристиками. В таком случае существует риск получить недостоверные оценки.

2. Методология

2.1. Проблема идентификации гетерогенности ответов при самооценке здоровья

Опишем возникающую проблему формальнее. Обозначим латентную переменную, описывающую состояние здоровья респондента с номером i , через H_i^* . Предполагается, что она ненаблюдаема и зависит от характеристик респондента X как

$$H_i^* = \beta_0 + X_i' \beta + \varepsilon_i, \quad \varepsilon_i \sim NID(0, \sigma^2). \quad (1)$$

Во время опроса респондент соотносит свое состояние здоровья с набором пороговых точек $-\infty = \tau_0 < \tau_1 < \dots < \tau_k = \infty$, и для оценки здоровья H_i выбирает один из вариантов ответов $k = 1, \dots, K$ согласно правилу

$$H_i = k, \quad \text{если } \tau_{k-1} \leq H_i^* < \tau_k. \quad (2)$$

В отличие от H_i^* , величина H_i наблюдаема. Если пороговые точки у всех респондентов одинаковы, то имеем стандартную модель упорядоченного выбора. Для идентифицируемости вводят нормализацию масштаба ($\sigma^2 = 1$) и положения ($\tau_1 = 0$ или $\beta_0 = 0$).

В общем случае наборы пороговых точек у респондентов могут отличаться. Пусть для респондента с номером i это есть набор $-\infty = \tau_0^{(i)} < \tau_1^{(i)} < \dots < \tau_k^{(i)} = \infty$. Предположим, что у респондентов из групп 1 и 2 уровень здоровья H^* одинаков, но $\tau_1^{(2)} < H^* < \tau_1^{(1)} < \tau_2^{(2)}$. Тогда респондент из группы 1 для оценки своих проблем выберет $k = 1$, а респондент из группы 2 выберет $k = 2$. Основываясь только на самооценках, делается вывод, что проблем у респондента

из первой группы меньше. Однако данный вывод может оказаться ошибочным, поскольку шкалы ответов у респондентов из 1-й и 2-й групп разные.

К сожалению, если истинное здоровье и пороговые точки шкалы ответов определяются одними и теми же переменными, невозможно различить связь этих переменных с истинным здоровьем и смещением пороговых точек. Действительно, гетерогенность ответов при самооценке здоровья можно моделировать как

$$\tau_k^{(i)} = \gamma_k^0 + X_i' \gamma_k, \quad (3)$$

тогда вероятность получить категорию k в ответ на вопрос о самооценке здоровья равна

$$P(H_i = k | X_i) = F\left(\frac{\gamma_k^0 + X_i' \gamma_k - \beta_0 - X_i' \beta}{\sigma}\right) - F\left(\frac{\gamma_{k-1}^0 + X_i' \gamma_{k-1} - \beta_0 - X_i' \beta}{\sigma}\right), \quad (4)$$

где F — функция распределения стандартного нормального закона. Ясно, что нельзя одновременно идентифицировать γ и β , если только наборы объясняющих переменных для самооценки здоровья и пороговых точек категорий ответов не различаются. Следовательно, если нет непосредственной информации об истинном здоровье H^* респондентов или об их пороговых точках $\tau_k^{(i)}$, то нет и возможности узнать из ответов респондентов, действительно ли у них разные проблемы, или же присутствует гетерогенность ответов.

В литературе существует как минимум два подхода к выявлению гетерогенности ответов при самооценке здоровья. Первый из них состоит в использовании объективных мер здоровья в качестве прокси-индикаторов истинного состояния здоровья (Lindeboom, Doorslaer, 2004). При этом используется предположение о том, что объективные индикаторы здоровья фиксируют все различия в истинном здоровье, а все оставшиеся различия в самооценке здоровья объясняются формированием мнения о состоянии здоровья, зависящего от характеристик индивида. Именно такой подход используется в работе (Канева, Байдин, 2018). Его недостаток состоит в том, что индикаторы здоровья, используемые в качестве объективных, также могут быть смещены. Индивиды с низким доходом могут не иметь диагностированных заболеваний в силу возможного ограничения доступа к медицинским услугам, меньшего объема диагностических процедур и, следовательно, менее вероятной постановки правильного диагноза (Burström, 2002; Newacheck et al., 2003). Для более образованных пациентов информация о симптомах и методах диагностики заболеваний может быть доступнее и понятнее, вследствие чего они лучше осведомлены об имеющихся заболеваниях (Terraneo, 2015). Респонденты также могут занижать число хронических заболеваний из-за опасений по поводу дискриминации на рынке труда (Gupta, Jürges, 2012). С другой стороны, согласно «гипотезе оправдывания», неработающие респонденты могут преувеличивать свои проблемы со здоровьем. В исследовании (Baker et al., 2004) было выявлено, что для многих заболеваний более половины респондентов с положительным диагнозом врача в медицинской документации не сообщили о наличии заболевания при проведении обследования, и наоборот, часть респондентов, заявивших о наличии заболевания, не имели соответствующих записей в медицинской документации.

Второй подход состоит в использовании так называемых якорных виньеток. Этот метод подробно описывается в следующем разделе.

2.2. Якорные виньетки: описание инструмента

Якорная виньетка — это набор сценариев, каждый из которых описывает специфическую ситуацию гипотетического индивида, относящуюся к измеряемому посредством самооценки феномену. Метод якорных виньеток был предложен в работе (King et al., 2004) как одно из решений проблемы измерения сложных понятий в опросах, связанной с тем, что респонденты по-разному интерпретируют одинаковые вопросы. Применимость метода к измерениям здоровья впервые тестировалась в работе (Murray et al., 2003). Якорные виньетки были использованы в исследованиях в различных тематических областях, включая здоровье (Bago d’Uva et al., 2008; Rossouw et al., 2018), трудоспособность (Karpeyn et al., 2007), качество системы здравоохранения (Valentine et al., 2015), удовлетворенность жизнью (Angelini et al., 2014). Методологическую литературу, посвященную данному методу, можно разделить на два направления. Одно из них связано с дизайном и методами сбора виньеток, включая предварительное тестирование опросного инструментария для снижения нагрузки на респондентов (King, Wand, 2007), порядок следования вопросов в анкете для снижения ошибок измерения (Hopkins, King, 2010), соответствие пола и возраста респондента и описываемого виньеткой гипотетического индивида (Grol-Prokopczyk, 2014). Другое направление исследует статистические свойства используемых в рамках данного метода моделей (Rice et al., 2011; Van Soest, Vonkova, 2014) и, в частности, тестирует выполнимость предпосылок, лежащих в основе метода (Bago d’Uva et al., 2011; Gupta et al., 2010; Peracchi, Rossetti, 2013; Van Soest et al., 2011).

Виньетки являются описаниями специфических состояний здоровья. Их используют для фиксации уровней истинного здоровья. Все систематические изменения рейтингов виньеток в зависимости от отдельных характеристик респондентов объясняются систематической разницей в формировании мнения о состоянии здоровья, т. е. различиями в их пороговых точках $\tau_k^{(i)}$.

Продemonстрируем метод на условном примере. Дадим респондентам в случайном порядке описание разного уровня проблем с концентрацией и памятью гипотетических людей.

V_1 . Александр очень быстро учится новым навыкам в своей работе. Он может удерживать внимание на своих рабочих заданиях в течение длительных непрерывных периодов времени. Он легко запоминает имена людей, адреса, номера телефонов, а также помнит детали событий, которые произошли несколько лет назад.

V_2 . Ирина может сосредоточиться во время просмотра телевизора, чтения журнала или игры в карты или шахматы. Она может понять новые комбинации в этих играх или их вариации с небольшими усилиями. Раз в неделю она забывает, где находятся ее ключи или очки, но находит их в течение пяти минут.

V_3 . Мария может совершать прогулки в окрестности своего дома и всегда находит путь домой, знает, где хранятся ее собственные вещи, но ей тяжело вспомнить, как добраться до места, которое она посещала только один или два раза. Ей интересно узнавать новые рецепты, но она часто обнаруживает, что делает ошибки, и должна перечитывать рецепт несколько раз, прежде чем сможет воспользоваться им.

V_4 . Иван не может сосредоточиться более чем на 15 минут и с трудом обращает внимание на то, что ему говорят. Всякий раз, когда он начинает какое-либо задание, он не справляется с ним и часто забывает, что он делает. Он может запомнить имена людей, которых встречает, но с трудом может дойти до магазина.

V_5 . Наталья не узнает даже близких родственников, не может найти дорогу домой, когда покидает дом без сопровождения. Даже с подсказками она не может вспомнить события или узнать родственников. Она не может приобретать новые знания, даже простые инструкции приводят ее в замешательство.

Описания $V_1, \dots, V_5$ называются якорными виньетками, будем обозначать их V_j .

Затем респондентов просят оценить, насколько трудно этим гипотетическим людям концентрироваться и запоминать вещи по шкале: «нет проблем», «незначительные», «умеренные», «сильные», «экстремальные». Заметим, что независимо от того, к какой группе относится респондент, оценивающий проблемы со здоровьем гипотетического индивида, сам гипотетический индивид, описанный виньеткой V_j , имеет один и тот же уровень истинного здоровья. Таким образом, создается «якорь», который позволяет выявить наличие или отсутствие систематической разницы в оценке респондентами того или иного уровня проблем со здоровьем в зависимости от характеристик респондента. Также метод якорных виньеток позволяет перевести самооценки этих групп в единую шкалу ответов, очищенную от систематического смещения, если оно обнаружено.

Для того чтобы метод якорных виньеток работал корректно, необходимо выполнение двух условий. Во-первых, должно выполняться так называемое условие эквивалентности виньетки. Это означает, что все респонденты понимают, что здоровье ухудшается от V_1 к V_5 . Во-вторых, респонденты должны одинаково использовать шкалы ответов, когда оценивают виньетки, и когда оценивают сами себя. Это условие называется согласованностью ответов. То есть пороговые точки, используемые индивидами для оценки своего здоровья и оценок виньеток, должны совпадать.

Разработчики виньеток стараются сконструировать их так, чтобы уменьшить возможное нарушение условий эквивалентности и согласованности. Для выполнения эквивалентности состояние гипотетических индивидов описывается как можно более конкретно, детализировано, без использования общих слов. Например, вместо «длинная дистанция» рекомендуется использовать «5 километров», вместо «не ориентируется в пространстве» — «не может найти близлежащий знакомый магазин», и т. д. Недостаточно конкретные и неподробные тексты виньеток служат причиной того, что респонденты могут начать вкладывать в них свои представления о том или ином состоянии здоровья (Van Soest et al., 2011). Для согласованности виньеток респонденты получают инструкции рассматривать и оценивать гипотетических персонажей виньеток, как самих себя (World Health Organization, 2006). Имена в виньетках выбираются из обычных женских/мужских имен с учетом страновой специфики. При тестировании данных предположений были получены противоречивые результаты. Некоторые исследования уверенно подтверждают их выполнение (Rice et al., 2011; van Soest et al., 2011), тогда как в других найдены нарушения (Gupta et al., 2010; Peracchi, Rossetti, 2013).

Якорные виньетки содержатся во многих обследованиях населения, в том числе во Всемирном исследовании здоровья населения (World Health Survey, WHS), Обследовании здоровья, старения и выхода на пенсию в Европе (Survey of Health, Ageing and Retirement in Europe, SHARE), Английском лонгитюдном исследовании старения (English Longitudinal Study of Ageing, ELSA), Исследовании здоровья и выхода на пенсию (Health and Retirement Study, HRS), Китайском лонгитюдном исследовании старения и выхода на пенсию (Chinese Health and Retirement Longitudinal Study, CHARLS) и Исследовании глобального старения и здоровья взрослого населения (Study on Global AGEing and Adult Health, SAGE).

Для анализа якорных виньеток могут применяться как параметрические, так и непараметрические методы (King et al., 2004). В настоящей статье будет использоваться параметрический подход, а именно, иерархическая модель упорядоченного выбора (hierarchical ordered probit, НОРИТ).

Данное исследование продолжает работу (Кучаков, Раскина, 2017), посвященную сравнению самооценок здоровья пожилого населения России и Китая, в которой использовался этот же метод. Новизна настоящей статьи заключается в следующем: а) описана и использована одношаговая процедура оценивания иерархической модели упорядоченного выбора; б) приведены результаты теста Вальда о наличии гетерогенности ответов; в) проведенный анализ включает 7 аспектов здоровья (14 индикаторов) в данных ВОЗ; г) оценка детерминант здоровья происходит в рамках страновой модели, что позволяет выявить эффект гетерогенности ответов отдельно для жителей России; д) представлен комплексный обзор методологической литературы по проблеме идентификации и коррекции самооценок респондентов с помощью метода якорных виньеток.

2.3. Иерархическая модель упорядоченного выбора НОРИТ

Модель является иерархической, поскольку состоит из двух последовательных частей: «уравнения якорных виньеток», позволяющего моделировать пороговые точки как функцию определяющих их характеристик индивида, и «уравнения здоровья», в котором самооценка здоровья сопоставляется с характеристиками индивида после коррекции на различия в восприятии уровня здоровья, полученными в первой части модели.

Воспринимаемый i -м респондентом уровень здоровья виньетки V_j равен

$$HV_{ij}^* = \alpha_j + u_{ij}, \quad u_{ij} \sim NID(0,1). \quad (5)$$

Для идентифицируемости вводят нормализацию масштаба $\alpha_1 = 0$. Во время опроса респондент с номером i соотносит состояние здоровья с набором своих пороговых точек⁶ $-\infty = \mu_0^{(i)} < \mu_1^{(i)} < \dots < \mu_5^{(i)} = \infty$, и для оценки здоровья HV_{ij} в виньетке V_j выбирает один из вариантов ответов $k = 1, \dots, 5$ согласно правилу:

$$HV_{ij} = k, \quad \text{если } \mu_{k-1}^{(i)} \leq HV_{ij}^* < \mu_k^{(i)}. \quad (6)$$

При этом пороговые точки могут зависеть от характеристик респондента:

$$\mu_k^{(i)} = \gamma_k^0 + Z_i' \gamma_k, \quad k = 1, \dots, 4 \quad (7)$$

(наборы регрессоров Z из (7) и X из (3) могут пересекаться, в данном исследовании они совпадают).

Используя упорядоченную пробит-модель, мы можем оценить γ и μ . Постоянные пороговые точки означают отсутствие гетерогенности ответов при самооценке здоровья, поэтому тестирование заключается в проверке гипотезы о незначимости компонент γ_k с помощью теста Вальда. Если эта гипотеза отвергается, то для коррекции смещения пороговых точек τ_k в «уравнении здоровья» (2) они заменяются оценками $\mu_k^{(i)}$, полученными в «уравнении якорных виньеток» (5)–(7). Оценивание модели НОРИТ также может производиться

⁶ В наших данных число ответов $K = 5$.

в один шаг, что эффективнее последовательного оценивания «уравнения якорных виньеток» и «уравнения здоровья», см. (Jones et al., 2012, p. 92). Логарифмическая функция правдоподобия при таком способе равна сумме двух компонент — самооценки здоровья и оценки описанного виньетками уровня здоровья:

$$\ln L = \sum_{i=1}^n \sum_{k=1}^5 h_{ik} \ln P_{ik} + \sum_{i=1}^n \sum_{j=1}^5 \sum_{k=1}^5 h_{ijk}^v \ln P_{ijk}, \quad (8)$$

где n — число респондентов; P_{ik} — вероятность получить категорию k в ответ на вопрос о самооценке здоровья респондента i , равная

$$P_{ik} = P(H_i = k | X_i) = F\left(\frac{\mu_k^{(i)} - \beta_0 - X_i' \beta}{\sigma}\right) - F\left(\frac{\mu_{k-1}^{(i)} - \beta_0 - X_i' \beta}{\sigma}\right);$$

h_{ik} — бинарная переменная, равная 1, если $H_i = k$, и 0 иначе; P_{ijk} — вероятность получить категорию k в ответ на вопрос об оценке здоровья в виньетке V_j для индивида i , равная $P_{ijk} = P(HV_{ij} = k | X_i) = F(\mu_k^{(i)} - \alpha_j) - F(\mu_{k-1}^{(i)} - \alpha_j)$; h_{ijk}^v — бинарная переменная, равная 1, если $HV_{ij} = k$, и 0 иначе.

Работа (Jones et al., 2012, p. 92–98) содержит пример применения модели НОРИТ для коррекции гетерогенности ответов самооценок здоровья и программный скрипт в пакете Stata для оценивания этой модели, который был использован в настоящем исследовании.

3. Данные и переменные

В статье используется репрезентативная выборка населения России 50 лет и старше из первой волны Исследования ВОЗ по проблемам глобального старения и здоровья взрослых людей (SAGE)⁷. Первая волна SAGE проходила в 2007–2010 гг. В настоящее время итоги второй (2014–2015 гг.) и третьей (2018 г.) волн недоступны для публичного использования. Как сообщается на сайте ВОЗ, получение российских данных находится под вопросом из-за изменения российского законодательства о персональных данных — команда ВОЗ ведет переговоры об их получении. Таким образом, данные первой волны — единственные доступные к настоящему времени. Возможно, данные последующих волн не будут доступны в обозримом будущем.

Для построения выборки использовался типичный дизайн для общенациональных опросов — вероятностная сложная многоуровневая выборка с выделением страт, формированием кластеров с присвоенными весами и последующим отбором, а также проведением итоговой постстратификации. В данной работе постстратификационные веса не использовались.

⁷ Документацию к набору данных первой волны российской выборки SAGE, включая методы построения выборки и сбора данных, опросные анкеты исследования, описание переменных, можно найти по адресу <https://apps.who.int/healthinfo/systems/surveydata/index.php/catalog/68>. Данные доступны для скачивания после запроса, который можно осуществить по приведенной выше ссылке.

Из выборки удалены респонденты, которые, по мнению интервьюеров, имеют когнитивные нарушения, и вместо них на анкеты отвечали их близкие (прокси-респонденты), а также респонденты младше 50 лет, т. к. эта группа не является целевой для SAGE и выборка младше 50 лет не является репрезентативной. Также в анализ не включены индивиды старше 80 лет из-за того, что эта группа респондентов малочисленна. Заметим, что анализ проводился еще двумя способами — без ограничения выборки по возрасту сверху и с ограничением 90 годами, при этом результаты качественно не менялись.

Центральное место в данной работе занимают вопросы о самооценке здоровья по восьми его аспектам (мобильность⁸, эмоциональное состояние, боль и скованность в теле, межличностные взаимодействия, зрение, сон и энергия, когнитивные способности, забота о себе) и связанные с ними якорные виньетки. Респондентов просят оценить проблемы, с которыми они сталкиваются в каждом из этих аспектов здоровья, используя вопрос «В целом за последние 30 дней, насколько большие проблемы у вас возникли с [аспект здоровья]?» по следующей шкале: 1 — нет, 2 — легкие, 3 — умеренные, 4 — тяжелые или 5 — экстремальные.

Анкета содержит два вопроса по двум конкретным индикаторам каждого аспекта здоровья (всего 8 аспектов, 16 индикаторов здоровья). Например, в аспекте когнитивных способностей респондента спрашивают, насколько ему/ей было трудно концентрироваться или запоминать вещи (индикатор «концентрация и память» аспекта «когнитивные способности») и при изучении новой задачи (индикатор «обучение» аспекта «когнитивные способности») за последние 30 дней. Используются все аспекты и составляющие их индикаторы здоровья, за исключением аспекта «зрение» и его индикаторов «близорукость» и «дальнозоркость», будем далее для простоты называть их 14-ю индикаторами здоровья. Распределение ответов респондентов на вопрос о проблемах для этих индикаторов здоровья приведены в табл. 1.

В таблицах 2, 3 и 4 в качестве иллюстрации различий в самооценке здоровья в зависимости от социально-демографических характеристик индивидов представлено распределение ответов для индикатора «обучение новому» в зависимости от уровня образования, для индикатора «сон» в зависимости от пола и для индикатора «боль» в зависимости от возраста, соответственно.

Анкета также содержит пять виньеток для каждого из восьми аспектов здоровья. Необходимость отвечать на множество однотипных вопросов приводит к эмоциональной усталости респондента. Для предотвращения отказов от участия в опросе вся выборка случайным образом делится на четыре равные подвыборки, в каждой из которых предлагается оценить виньетки в двух из восьми аспектов здоровья. Таким образом, каждый из двух аспектов (четыре индикатора) оценивается четвертью всей выборки. Из выборки удаляются наблюдения, для которых отсутствует ответ хотя бы в одной из пяти виньеток для каждого отдельного индикатора здоровья.

Если эквивалентность виньеток выполняется, виньетка V_1 (описывающая отсутствие проблем в каждом аспекте здоровья) должна пониматься всеми респондентами как лучший уровень здоровья, чем виньетка V_2 (описывающая незначительные проблемы в каждом аспекте здоровья) и так далее, т. е. $V_1 < V_2 < V_3 < V_4 < V_5$. Если у какого-либо респондента этот порядок нарушается, то он либо не понимает вопроса, либо трактует концепцию состояния здоровья каким-либо иным образом, чем составители анкеты, и эквивалентность виньеток нарушается. «Связанные ранги», т. е. ситуацию, когда респондент ранжирует две соседние

⁸ Способность перемещаться в пространстве.

Таблица 1. Распределение ответов о проблемах для индикаторов здоровья по степени тяжести

Аспект / индикатор здоровья	% ответов					Число ответов
	1 нет	2 легкие	3 умеренные	4 тяжелые	5 экстремальные	
<i>Физическая мобильность</i>						
Мобильность	43.16	25.77	19.82	9.83	1.42	3531
Повседневная активность	19.99	30.12	27.15	16.96	5.78	3496
<i>Эмоциональное состояние</i>						
Депрессивное настроение	59.90	23.66	12.18	3.72	0.55	3466
Тревожность	49.94	31.87	13.67	3.74	0.78	3474
<i>Боль и дискомфорт в теле</i>						
Боль	36.87	29.13	22.26	10.29	1.45	3526
Скованность	35.90	30.05	22.02	10.67	1.36	3524
<i>Межличностные взаимодействия</i>						
Взаимоотношения	74.06	16.92	6.06	1.85	1.11	3516
Конфликты	66.07	24.84	6.70	1.82	0.57	3507
<i>Сон и энергия</i>						
Засыпание	28.99	36.78	22.49	10.66	1.09	3491
Энергия после пробуждения	25.77	39.61	25.09	8.55	0.98	3484
<i>Когнитивные способности</i>						
Память	47.72	32.94	14.75	4.43	0.17	3525
Обучение новому	49.39	30.24	13.97	5.86	0.54	3515
<i>Забота о себе</i>						
Самообслуживание	72.64	14.83	8.82	2.89	0.82	3527
Внешний вид	81.44	10.60	6.46	1.08	0.43	3529

Таблица 2. Распределение ответов о проблемах для индикатора «обучение новому» в зависимости от уровня образования, в процентах

Шкала проблем	Уровень образования		
	Основное общее образование или ниже	Среднее (полное) общее образование	Профессиональное образование (среднее профессиональное образование и выше)
1 — нет	32.90	52.04	64.84
2 — легкие	31.91	31.76	24.62
3 — умеренные	22.84	11.64	7.93
4 — тяжелые	11.26	4.30	2.19
5 — экстремальные	1.09	0.27	0.41

виньетки одинаково (например, $V_1 < V_2 = V_3 < V_4 < V_5$), трактуют как не нарушающие порядок виньеток (Murray et al., 2003).

В таблице 5 представлен процент респондентов, которые в своих ответах не нарушают порядок ранжирования виньеток, а также тех, кто нарушает этот порядок только для одной пары виньеток. В работе (Murray et al., 2003) такой порядок называется почти последовательным (near consistent), и такие случаи не рассматриваются как «брак». Доля респондентов,

Таблица 3. Распределение ответов о проблемах для индикатора «сон» в зависимости от пола, в процентах

Шкала проблем	Пол	
	Мужчины	Женщины
1 — нет	38.67	23.47
2 — легкие	35.75	37.37
3 — умеренные	18.15	24.96
4 — тяжелые	6.95	12.77
5 — экстремальные	0.47	1.44

Таблица 4. Распределение ответов о проблемах для индикатора «боль» в зависимости от возраста, в процентах

Шкала проблем	Возраст, в годах					
	50–54	55–59	60–64	65–69	70–74	75–79
1 — нет	53.67	45.02	40.08	29.48	23.42	18.64
2 — легкие	26.28	28.38	33.20	32.19	27.29	28.86
3 — умеренные	13.83	18.83	17.29	24.95	33.10	30.23
4 — тяжелые	5.67	6.82	7.86	11.93	13.73	20.00
5 — экстремальные	0.55	0.95	1.57	1.45	2.46	2.27

Таблица 5. Число и доля респондентов, нарушающих порядок ранжирования виньеток

Аспект / индикатор здоровья	Нет нарушения порядка ранжирования		Ранжирование нарушено только для одной пары виньеток	
	Человек	%	Человек	%
<i>Физическая мобильность</i>				
Мобильность	671	79.69	835	99.17
Повседневная активность	633	75.72	821	98.21
<i>Эмоциональное состояние</i>				
Депрессивное настроение	693	82.50	825	98.21
Тревожность	663	79.21	825	98.56
<i>Боль и дискомфорт в теле</i>				
Боль	650	79.75	797	97.79
Скованность	654	80.64	791	97.53
<i>Межличностные взаимодействия</i>				
Взаимоотношения	524	63.90	782	95.36
Конфликты	526	64.70	783	96.31
<i>Сон и энергия</i>				
Засыпание	382	47.51	766	95.27
Энергия после пробуждения	416	51.87	763	95.14
<i>Когнитивные способности</i>				
Память	699	84.42	819	98.91
Обучение новому	703	85.32	816	99.03
<i>Забота о себе</i>				
Самообслуживание	552	66.43	810	97.48
Внешний вид	578	69.98	802	97.10

корректно ранжирующих виньетки, для восьми индикаторов выше 75%, однако есть индикаторы, где она значительно ниже. Так, для индикатора «сон» она опускается ниже 50%, у половины выборки порядок ранжирования виньеток нарушен. Доля респондентов, нарушивших порядок ранжирования не более одного раза, выше 95% во всех случаях. Хотя в (Murray et al., 2003) рекомендуется не браковать такие случаи, в настоящей статье применяется наиболее жесткий критерий и из дальнейшего анализа удаляются все респонденты, нарушившие порядок ранжирования (связанные ранги при этом возможны) хотя бы один раз. Наконец, для того чтобы модель упорядоченного выбора и НОРІТ модель оценивались по одинаковым выборкам, удаляются наблюдения, в которых пропущен ответ для оценки своего здоровья для каждого индикатора здоровья.

Для измерения экономического благосостояния используются квантильные группы индекса перманентного дохода, рассчитанного на основании данных о располагаемых ресурсах, наличии товаров длительного пользования, жилищных условий по модели, предложенной в работе (Ferguson et al., 2003) в таком виде, как они находятся в данных SAGE. Перманентный доход лучше отражает уровень благосостояния человека, чем текущий, особенно для людей пожилого возраста, текущие доходы которых, как правило, невелики.

Уровень образования измеряется наивысшим достигнутым его уровнем. Вопросник SAGE содержит шестиуровневую градацию образовательного уровня, однако среди них встречаются очень немногочисленные категории, поэтому некоторые образовательные уровни объединяются и перекодируются в трехуровневую градацию. Авторы используют общепринятый в литературе набор переменных: возраст, пол, семейный статус и место проживания (город или село), а также менее подробную классификацию семейного положения, чем содержится в опроснике SAGE — по тем же причинам, что и для образования.

Описательные статистики переменных, которые используются для выявления и коррекции самооценки здоровья, приведены в табл. 6.

4. Результаты

В данном исследовании были поставлены две задачи:

- 1) выяснить, есть ли гетерогенность ответов при самооценке здоровья россиян; если да, то для каких аспектов здоровья и какие социально-экономические факторы ее вызывают;
- 2) оценить связь социально-экономических факторов и здоровья россиян, скорректированную на гетерогенность ответов.

Рассмотрим для примера результаты оценивания модели НОРІТ для индикатора здоровья «боль в теле», приведенные в табл. 7. Обсудим фиктивную переменную для первой квантильной группы индекса перманентного дохода. Коэффициент для пороговой точки 3, отделяющий категорию «умеренные проблемы» от «тяжелые проблемы» значимо отличен от нуля и положителен. Это означает, что у людей из первой квантильной группы, самой малообеспеченной, эта пороговая точка смещена вправо по сравнению с людьми из наиболее обеспеченной пятой (базовой) квантильной группы. При одном и том же истинном уровне испытываемых проблем с болями в теле менее обеспеченные люди будут скорее оценивать его как умеренный, а более обеспеченные — как тяжелый.

Для разных индикаторов здоровья смещение пороговых точек может зависеть от разных факторов. Тестирование отсутствия гетерогенности ответов при самооценке здоровья

Таблица 6. Описательные статистики

Переменные в наборе данных		Перекодированные переменные	
<i>Пол</i>			
Мужчины (%)	36.31		
Женщины (%)	63.69		
<i>Возраст</i>			
Среднее (стандартное отклонение)	63.10 (8.59)		
<i>Образование, высший уровень</i>		<i>Образование, высший уровень</i>	
Не закончил начальную школу (%)	1.69	Незаконченное среднее и меньше (%)	26.40
Начальная школа (%)	6.19	Законченное среднее (%)	52.71
Незаконченное среднее (%)	18.61	Колледж/университет и выше (%)	20.89
Законченное среднее о (%)	52.71		
Колледж/университет (%)	20.57		
Аспирантура	0.31		
<i>Квантили перманентного дохода</i>			
Q1 (беднейшие)	17.63		
Q2	19.09		
Q3	19.77		
Q4	20.76		
Q5	22.76		
<i>Семейное положение</i>		<i>Семейное положение</i>	
Никогда не был в браке (%)	2.08	Живет раздельно / Разведен /	40.44
В настоящее время в браке (%)	56.46	Вдовствует / Никогда не был в браке (%)	
Проживает с партнером (%)	3.11	В настоящее время в браке /	59.56
Живет раздельно / Разведен (%)	8.93	Проживает с партнером (%)	
Вдовствует (%)	28.71		
<i>Место жительства</i>			
Село (%)	24.73		
Город (%)	75.27		

Примечание. В таблице приведены описательные статистики для выборки после удаления из нее данных о людях старше 80 лет, но до удаления наблюдений, в которых пропущены ответы для оценки виньеток и самооценки здоровья для каждого индикатора. В подвыборках (14 штук), для которых присутствуют наблюдения по всему набору виньеток и самооценки здоровья для каждого отдельного индикатора, статистики незначительно изменяются. Для экономии места описательные статистики по каждой подвыборке не приводятся.

по каждой переменной, включенной в вектор Z (см. уравнение (7)), может быть выполнено путем проверки значимости соответствующих компонент вектора γ . Таблица 8 содержит результаты проверки подобных нулевых гипотез тестом Вальда для всех переменных, включенных в уравнение (7).

Согласно результатам тестов, смещение пороговых точек, зависящее от характеристик респондента, присутствует для всех аспектов здоровья. Смещение пороговых точек связано с перманентным доходом для почти всех аспектов и индикаторов здоровья, за исключением аспекта «забота о себе» и индикатора «повседневная активность». Люди с разным уровнем образования по-разному определяют пороговые точки для аспектов «эмоциональное состояние» и «забота о себе», а также индикаторов «мобильность», «боль», «запоминание». Смещение пороговых точек связано с возрастом для аспектов здоровья «межличностные

Таблица 7. Результаты оценивания модели НОРІТ, коэффициенты при пороговых точках, индикатор здоровья — боль в теле

Переменные	Уравнение собственного здоровья	Пороговые точки			
		1: нет vs. легкие	2: легкие vs. умеренные	3: умеренные vs. тяжелые	4: тяжелые vs. экстремальные
Q1 перманентного дохода (беднейшие)	0.849*** (0.290)	0.327** (0.140)	0.049 (0.105)	0.188* (0.112)	−0.027 (0.131)
Q2 перманентного дохода	0.313 (0.283)	0.059 (0.143)	0.066 (0.103)	0.137 (0.110)	0.167 (0.128)
Q3 перманентного дохода	0.415 (0.275)	0.070 (0.136)	−0.047 (0.099)	0.004 (0.106)	0.062 (0.123)
Q4 перманентного дохода	0.420 (0.268)	0.472*** (0.125)	0.093 (0.095)	−0.001 (0.103)	0.011 (0.118)
Незаконченное среднее и меньше	0.577** (0.291)	−0.351** (0.141)	−0.092 (0.105)	0.088 (0.113)	−0.030 (0.130)
Законченное среднее	0.119 (0.242)	−0.288*** (0.110)	−0.201** (0.084)	−0.081 (0.091)	0.182* (0.105)
Возраст	0.062*** (0.012)	−0.008 (0.006)	0.003 (0.004)	0.000 (0.005)	0.003 (0.005)
Пол (1 — женский, 0 — мужской)	0.546*** (0.194)	−0.121 (0.092)	−0.091 (0.070)	−0.043 (0.075)	0.032 (0.087)
Место жительства (1 — город, 0 — село)	−0.087 (0.215)	−0.080 (0.103)	0.083 (0.077)	−0.041 (0.084)	−0.202** (0.098)
Семейный статус (1 — проживает с партнером, 0 — проживает один)	0.025 (0.203)	0.002 (0.010)	−0.047 (0.074)	0.142* (0.080)	0.185** (0.093)
Число наблюдений		645			
LR-тест		76.6***			

Примечание. *, **, *** соответствуют 10, 5 и 1%-ному уровню значимости, соответственно. В скобках представлены стандартные ошибки. LR-тест — статистика отношения правдоподобия для проверки нулевой гипотезы о совместной незначимости всех объясняющих переменных в модели.

взаимодействия», «когнитивные способности», «забота о себе», а также индикаторами «мобильность» и «тревожность». Вопреки довольно распространенному мнению, что женщины и мужчины сильно отличаются при формировании мнения о своем здоровье, пол смещает пороговые точки только для индикатора «взаимоотношения» (на 10%-ном уровне значимости). Городские и сельские жители формируют различные пороговые точки для трех аспектов («сон и энергия», «межличностные взаимодействия», «забота о себе») и для индикатора «повседневная активность». Семейный статус связан со смещением пороговых точек для аспекта «боль и дискомфорт в теле» и индикатора «поддержание внешнего вида».

В таблице А1 (в Приложении) приведены результаты оценивания упорядоченной пробит-модели и модели НОРІТ для самооценки здоровья по его 14 индикаторам. Жирным шрифтом в ней выделены коэффициенты, значимость которых (как минимум на 10%-ном уровне)

Таблица 8. Результаты теста на отсутствие гетерогенности ответов при самооценке здоровья

Аспект / индикатор здоровья	Переменные						
	All variables	Permanent income all Q's	Education all 3cat	Age	Female	Urban	Currently married / cohab
<i>Физическая мобильность</i>							
Мобильность	***	***	***	***	×	×	×
Повседневная активность	*	×	×	×	×	***	×
<i>Эмоциональное состояние</i>							
Депрессивное настроение	***	***	**	×	×	×	×
Тревожность	***	***	*	***	×	×	×
<i>Боль и дискомфорт в теле</i>							
Боль	***	***	***	×	×	×	*
Скованность	***	**	×	×	×	×	***
<i>Межличностные взаимодействия</i>							
Взаимоотношения	***	***	×	***	*	***	×
Конфликты	***	***	×	***	×	**	×
<i>Сон и энергия</i>							
Засыпание	***	***	×	×	×	**	×
Энергия после пробуждения	***	***	×	×	×	***	×
<i>Когнитивные способности</i>							
Память	***	***	***	***	×	×	×
Обучение новому	***	***	×	***	×	×	×
<i>Забота о себе</i>							
Самообслуживание	***	×	***	***	×	***	***
Внешний вид	***	×	***	***	×	***	×

Примечание. ***, **, * и × соответствуют 1, 5, 10% и >10%-ному уровню значимости для теста Вальда с нулевой гипотезой о том, что коэффициенты при переменных, указанных в заголовках столбцов, статистически незначимы в модели (7).

не совпадает для пробит-модели упорядоченного выбора и модели НОРІТ. Они встречаются в переменных перманентного дохода, уровня образования, семейного статуса и места жительства, и не наблюдаются в переменных возраста и пола.

Обсудим детерминанты здоровья, используя результаты модели НОРІТ, учитывающей смещение пороговых точек. Для восьми индикаторов в четырех аспектах («физическая мобильность», «эмоциональное состояние», «боль и дискомфорт в теле», «забота о себе»), наблюдается важность дохода. Женщины чаще сообщают о проблемах со здоровьем в трех аспектах — «эмоциональное состояние», «боль и дискомфорт в теле», «сон и энергия». Довольно ожидаемо, что возраст связан с низкой оценкой состояния здоровья для всех индикаторов, кроме конфликтов. Люди, имеющие партнера, реже указывают на проблемы для индикатора «конфликты», чем одиночки. Уровень образования важен при самооценке здоровья в двух аспектах («боль и дискомфорт в теле» и «когнитивные способности»), а также для индикаторов «взаимоотношения» и «самообслуживание».

5. Заключение

Представленная работа демонстрирует недостатки, возникающие в исследованиях оценок здоровья с использованием самооценок респондентов без учета гетерогенности ответов. Использование якорных виньеток позволяет заключить следующее.

Во-первых, во всех 7 аспектах и составляющих их 14 индикаторов здоровья наблюдается гетерогенность ответов при самооценке здоровья. Перманентный доход связан со смещением пороговых точек для 11 индикаторов здоровья, образование и место проживания — для 7 индикаторов, возраст — для 8, семейный статус — для 3. Следует отметить, что широко распространенное мнение о существовании значимых различий между мужчинами и женщинами при оценивании своего здоровья не нашло подтверждения. Из всех 14 индикаторов здоровья пол смещает пороговые точки только для одного индикатора (взаимоотношения). Во-вторых, результаты регрессионного анализа с учетом гетерогенности ответов при самооценке здоровья и без нее отличаются. Перманентный доход, уровень образования, место проживания и семейный статус для некоторых индикаторов здоровья меняют свою значимость/незначимость в стандартной модели упорядоченного выбора и в модели НОРИТ. Таким образом, игнорирование гетерогенности в самооценках здоровья может стать причиной ненадежных выводов.

Использование опросных методов в качестве инструмента сбора данных о состоянии здоровья населения имеет ряд преимуществ, наиболее сильными из которых с точки зрения практической реализуемости является низкая цена и обилие получаемой информации. При этом в литературе, как уже было отмечено, существует множество вопросов к надежности и валидности самооценок здоровья. Вероятно, лучшим вариантом сбора данных являлось бы массовое медицинское обследование граждан, но по ряду причин оно малореалистично. Улучшение качества опросных данных за счет использования якорных виньеток представляется авторам способом, оставляющим сбор данных осуществимым. Именно поэтому можно рекомендовать использовать якорные винюетки при проведении опросов социально-экономической направленности.

Литература

Герри К. Д., Байдин В. М. (2017). Источники искажений в самооценке здоровья. *Фармакоэкономика. Современная фармакоэкономика и фармакоэпидемиология*, 10 (4), 31–36. DOI: 10.17749/2070-4909.2017.10.4.031-036.

Кучаков Р. К., Раскина Ю. В. (2017). Социально-экономический градиент самооценок здоровья пожилого населения России и Китая: результаты применения иерархических моделей упорядоченного выбора. *Финансы и бизнес*, 13 (4), 22–46.

Канева М. А., Байдин В. М. (2018). Гетерогенность ответов при самооценке здоровья россиян. *Прикладная эконометрика*, 3 (51), 102–125.

Пол П., Валтонен Х., Ковтун Н. В. (2019). Социально-экономическое неравенство населения в области здоровья в постсоветской России. *Народонаселение*, 22 (1), 61–78. DOI: 10.19181/1561-7785-2019-00005.

Хабибуллина А. Р., Александрова Е. А., Аистов А. В. (2021). Оценки неравенства в здоровье россиян. *Вопросы экономики*, 3, 117–138. DOI: 10.32609/0042-8736-2021-3-117-138.

Angelini V., Cavapozzi D., Corazzini L., Paccagnella O. (2014). Do Danes and Italians rate life satisfaction in the same way? Using vignettes to correct for individual-specific scale biases. *Oxford Bulletin of Economics and Statistics*, 76 (5), 643–666. DOI: 10.1111/obes.12039.

Bago d'Uva T., Doorslaer van E., Lindeboom M., O'Donnell O. (2008). Does reporting heterogeneity bias the measurement of health disparities? *Health Economics*, 17 (3), 351–375. DOI: 10.1002/hec.1269.

Bago d'Uva T., Lindeboom M., O'Donnell O., Doorslaer van E. (2011). Slipping anchor? Testing the vignettes approach to identification and correction of reporting heterogeneity. *Journal of Human Resources*, 46 (4), 875–906. DOI: 10.1353/jhr.2011.0005.

Baker M., Stabile M., Deri C. (2004). What do self-reported, objective, measures of health measure? *Journal of Human Resources*, 39 (4), 1067–1093. DOI: 10.2307/3559039.

Bambra C., Popham F. (2010). Worklessness and regional differences in the social gradient in general health: Evidence from the 2001 English census. *Health & Place*, 16 (5), 1014–1021. DOI: 10.1016/j.healthplace.2010.06.006.

Burström B. (2002). Increasing inequalities in health care utilisation across income groups in Sweden during the 1990s? *Health Policy*, 62 (2), 117–129. DOI: 10.1016/s0168-8510(02)00016-7.

Cutler D. M., Lleras-Muney A. (2006). Education and health: Evaluating theories and evidence. *Working Paper* 12352. National Bureau of Economic Research. DOI: 10.3386/w12352.

DeSalvo K. B., Bloser N., Reynolds K., He J., Muntner P. (2006). Mortality prediction with a single general self-rated health question. *Journal of General Internal Medicine*, 21 (3), 267–275. DOI: 10.1111/j.1525-1497.2005.00291.x.

Doorslaer van E., Koolman X. (2004). Explaining the differences in income-related health inequalities across European countries. *Health Economics*, 13 (7), 609–628. DOI: 10.1002/hec.918.

Eikemo T. A., Bambra C., Joyce K., Dahl E. (2008). Welfare state regimes and income-related health inequalities: A comparison of 23 European countries. *European Journal of Public Health*, 18 (6), 593–599. DOI: 10.1093/eurpub/ckn092.

Ferguson B. D., Tandon A., Gakidou E., Murray C. J. (2003). Estimating permanent income using indicator variables. In: C. J. L. Murray, D. B. Evans (eds.). *Health Systems Performance Assessment: Debates, Methods and Empiricism*, 747–760. Geneva: World Health Organization.

Grol-Prokopczyk H. (2014). Age and sex effects in anchoring vignette studies: Methodological and empirical contributions. *Survey Research Methods*, 8 (1), 1–17. DOI: 10.18148/srm/2014.v8i1.5133.

Gorshkov I., Zharinova A. (2016). The effect of education on health perception and health related behavior: Evidence from the Russian Federation. DOI: 10.2139/ssrn.2827370.

Groot W., Den Brink van H. M. (2007). The health effects of education. *Economics of Education Review*, 26 (2), 186–200. DOI: 10.1016/j.econedurev.2005.09.002.

Gupta N. D., Jürges H. (2012). Do workers underreport morbidity? The accuracy of self-reports of chronic conditions. *Social Science & Medicine*, 75 (9), 1589–1594. DOI: 10.1016/j.socscimed.2012.07.013.

Gupta N. D., Kristensen N., Pozzoli D. (2010). External validation of the use of vignettes in cross-country health studies. *Economic Modelling*, 27 (4), 854–865. DOI: 10.1016/j.econmod.2009.11.007.

Hopkins D. J., King G. (2010). Improving anchoring vignettes: Designing surveys to correct interpersonal incomparability. *Public Opinion Quarterly*, 74 (2), 201–222. DOI: 10.1093/poq/nfq011.

Idler E. L., Benyamini Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior*, 38 (1), 21–37. DOI: 10.2307/2955359.

Jones A. M., Rice N., Bago d'Uva T., Balia S. (2012). *Applied health economics*. Routledge. DOI: 10.4324/9780203102411.

Jürges H. (2007). True health vs response styles: Exploring cross-country differences in self-reported health. *Health Economics*, 16 (2), 163–178. DOI: 10.1002/hec.1134.

Jylhä M. (2009). What is self-rated health and why does it predict mortality? Towards a unified conceptual model. *Social Science & Medicine*, 69 (3), 307–316. DOI: 10.1016/j.socscimed.2009.05.013.

Jylhä M., Volpato S., Guralnik J. M. (2006). Self-rated health showed a graded association with frequently used biomarkers in a large population sample. *Journal of Clinical Epidemiology*, 59 (5), 465–471. DOI: 10.1016/j.jclinepi.2005.12.004.

Kapteyn A., Smith J. P., van Soest A. H. O. (2007). Vignettes and self-reports of work disability in the United States and the Netherlands. *American Economic Review*, 97 (1), 461–473. DOI: 10.1257/aer.97.1.461.

Kerkhofs M., Lindeboom M. (1995). Subjective health measures and state dependent reporting errors. *Health Economics*, 4 (3), 221–235. DOI: 10.1002/hec.4730040307.

King G., Murray C. J., Salomon J. A., Tandon A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review*, 98 (1), 191–207. DOI: 10.1017/S000305540400108X.

King G., Wand J. (2007). Comparing incomparable survey responses: Evaluating and selecting anchoring vignettes. *Political Analysis*, 15 (1), 46–66. DOI: 10.1093/pan/aml011.

Krieger N., Chen J. T., Waterman P. D., Rehkopf D. H., Subramanian S. V. (2003). Race/ethnicity, gender, and monitoring socioeconomic gradients in health: A comparison of area-based socioeconomic measures — the public health disparities geocoding project. *American Journal of Public Health*, 93 (10), 1655–1671. DOI: 10.2105/ajph.93.10.1655.

Lindeboom M., Doorslaer van E. (2004). Cut-point shift and index shift in self-reported health. *Journal of Health Economics*, 23 (6), 1083–1099. DOI: 10.1016/j.jhealeco.2004.01.002.

Lokshin M., Ravallion M. (2008). Testing for an economic gradient in health status using subjective data. *Health Economics*, 17 (11), 1237–1259. DOI: 10.1093/ije/30.3.600.

Manor O., Matthews S., Power C. (2001). Self-rated health and limiting longstanding illness: Interrelationships with morbidity in early adulthood. *International Journal of Epidemiology*, 30 (3), 600–607. DOI: 10.1093/ije/30.3.600.

Marmot M., Friel S., Bell R., Houweling T. A., Taylor S. (2008). Closing the gap in a generation: Health equity through action on the social determinants of health. *The Lancet*, 372 (9650), 1661–1669. DOI: 10.1016/S01406736(08)61690-6.

Martinson M. L. (2012). Income inequality in health at all ages: A comparison of the United States and England. *American Journal of Public Health*, 102 (11), 2049–2056. DOI: 10.2105/AJPH.2012.300929.

Miao J., Wu X. (2016). Urbanization, socioeconomic status and health disparity in China. *Health & Place*, 42, 87–95. DOI: 10.1016/j.healthplace.2016.09.008.

Molina T. (2017). Adjusting for heterogeneous response thresholds in cross-country comparisons of self-reported health. *The Journal of the Economics of Ageing*, 10, 1–20. DOI: 10.1016/j.jeoa.2017.08.002.

Montez J. K., Berkman L. F. (2013). Trends in the educational gradient of mortality among US adults aged 45 to 84 years: Bringing regional context into the explanation. *American Journal of Public Health*, 104 (1), 82–90. DOI: 10.2105/AJPH.2013.301526.

Murray C. J., Ozaltin E., Tandon A., Salomon J., Sadana R., Chatterji S. (2003). Empirical evaluation of the anchoring vignette approach in health surveys. In: C. J. L. Murray, D. B. Evans (eds.). *Health Systems Performance Assessment: Debates, Methods and Empiricism*, 369–399. Geneva: World Health Organization.

Newacheck P. W., Hung Y. Y., Jane Park M., Brindis C. D., Irwin Jr C. E. (2003). Disparities in adolescent health and health care: Does socioeconomic status matter? *Health Services Research*, 38 (5), 1235–1252. DOI: 10.1111/1475-6773.00174.

Nielsen A. B. S., Siersma V., Conradsen Hiort L., Drivsholm T., Kreiner S., Hollnagel H. (2008). Self-rated general health among 40-year-old Danes and its association with all-cause mortality at 10-, 20-, and 29 years follow-up. *Scandinavian Journal of Public Health*, 36 (1), 3–11. DOI: 10.1177/1403494807085242.

Oksuzyan A., Daňko M. J., Caputo J., Jasilionis D., Shkolnikov V. M. (2019). Is the story about sensitive women and stoical men true? Gender differences in health after adjustment for reporting behavior. *Social Science & Medicine*, 228, 41–50. DOI: 10.1016/j.socscimed.2019.03.002.

Paul P., Valtonen H. (2016). Inequalities in perceived health in the Russian Federation. 1994–2012. *BMC Public Health*, 16 (1), 1–12. DOI: 10.1186/s12889-016-2810-x.

Peracchi F., Rossetti C. (2013). The heterogeneous thresholds ordered response model: Identification and inference. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176 (3), 703–722. DOI: 10.1111/j.1467-985X.2012.01066.x.

Rice N., Robone S., Smith P. (2011). Analysis of the validity of the vignette approach to correct for heterogeneity in reporting health system responsiveness. *The European Journal of Health Economics*, 12 (2), 141–162. DOI: 10.1007/s10198-010-0235-5.

Riva M., Curtis S., Gauvin L., Fagg J. (2009). Unravelling the extent of inequalities in health across urban and rural areas: Evidence from a national sample in England. *Social Science & Medicine*, 68 (4), 654–663. DOI: 10.1016/j.socscimed.2008.11.024.

Rossouw L., Bago d’Uva T., Doorslaer van E. (2018). Poor health reporting? Using anchoring vignettes to uncover health disparities by wealth and race. *Demography*, 55 (5), 1935–1956. DOI: 10.1007/s13524-018-0709-0.

Van Soest A. H. O., Delaney L., Harmon C., Kapteyn A., Smith J. P. (2011). Validating the use of anchoring vignettes for the correction of response scale differences in subjective questions. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174 (3), 575–595. DOI: 10.1111/j.1467-985X.2011.00694.x.

Van Soest A. H. O., Vonkova H. (2014). Testing the specification of parametric models by using anchoring vignettes. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 177 (1), 115–133. DOI: 10.1111/j.1467-985X.2012.12000.x.

Stafford M., Cummins S., Macintyre S., Ellaway A., Marmot M. (2005). Gender differences in the associations between health and neighbourhood environment. *Social Science & Medicine*, 60 (8), 1681–1692. DOI: 10.1016/j.socscimed.2004.08.028.

Terraneo M. (2015). Inequities in health care utilization by people aged 50+: Evidence from 12 European countries. *Social Science & Medicine*, 126, 154–163. DOI: 10.1016/j.socscimed.2014.12.028.

Valentine N., Verdes-Tennant E., Bonsel G. (2015). Health systems’ responsiveness and reporting behaviour: Multilevel analysis of the influence of individual-level factors in 64 countries. *Social Science & Medicine*, 138, 152–160. DOI: 10.1016/j.socscimed.2015.04.022.

Scott N. W. et al. (2010). Differential item functioning (DIF) analyses of health-related quality of life instruments using logistic regression. *Health and Quality of Life Outcomes*, 8 (1), 1–9. DOI: 10.1186/1477-7525-8-81.

Schneider U., Pfarr C., Schneider B. S., Ulrich V. I. (2012). I feel good! Gender differences and reporting heterogeneity in self-assessed health. *The European Journal of Health Economics*, 13 (3), 251–265. DOI: 10.1007/s10198-011-0301-7.

World Health Organization. (2006). WHO SAGE survey manual: The WHO study on global ageing and adult health (SAGE). Geneva: World Health Organization.

Поступила в редакцию 07.09.2021;
принята в печать 17.05.2022.

Приложение

Таблица А1. Результаты оценивания модели самооценки здоровья: упорядоченный пробит и НОРИТ

Индикатор здоровья (n = число наблюдений)	Модель	Переменные						Пол (1 — женский, 0 — мужской)	Место жительства (1 — город, 0 — село)	Семейный статус (1 — проживает с партнером, 0 — проживает один)	Log-likelihood
		Q1 (беднейшие)	Q2	Q3	Q4	Незаконченное среднее и меньше	Законченное среднее				
Мобильность (n = 664)	oprobit	0.438*** (0.154)	0.285** (0.145)	0.204 (0.142)	0.099 (0.142)	0.197 (0.140)	0.098 (0.117)	0.124 (0.099)	-0.214** (0.107)	-0.171* (0.098)	-821.862
	НОРИТ	0.619** (0.275)	0.504* (0.259)	0.424* (0.253)	0.321 (0.253)	0.231 (0.248)	0.066 (0.208)	0.287 (0.176)	-0.392** (0.191)	-0.244 (0.174)	-3272.630
	oprobit	0.506*** (0.156)	0.283** (0.142)	0.166 (0.136)	0.078 (0.135)	0.063 (0.136)	0.046 (0.109)	0.079 (0.095)	-0.012 (0.109)	-0.100 (0.096)	-867.119
	НОРИТ	0.448** (0.220)	0.249 (0.201)	0.248 (0.194)	0.062 (0.192)	0.040 (0.193)	0.039 (0.156)	0.098 (0.134)	-0.189 (0.153)	-0.097 (0.136)	-3367.389
Депрессивное настроение (n = 683)	oprobit	0.551*** (0.163)	0.260* (0.156)	0.281* (0.151)	0.397*** (0.151)	-0.001 (0.144)	0.081 (0.119)	0.199* (0.103)	0.021 (0.113)	-0.036 (0.102)	-680.385
	НОРИТ	0.775** (0.306)	0.178 (0.292)	0.640** (0.280)	0.802*** (0.281)	0.187 (0.265)	0.137 (0.221)	0.324* (0.191)	0.050 (0.208)	-0.098 (0.188)	-3252.288
	oprobit	0.445*** (0.157)	0.144 (0.151)	0.199 (0.144)	0.270* (0.144)	-0.050 (0.140)	0.040 (0.116)	0.466*** (0.103)	-0.064 (0.113)	0.069 (0.102)	-703.142
	НОРИТ	0.559** (0.234)	0.160 (0.227)	0.439** (0.214)	0.610*** (0.215)	0.089 (0.208)	0.013 (0.173)	0.519*** (0.155)	-0.178 (0.167)	0.012 (0.151)	-3300.082
Боль (n = 645)	oprobit	0.339** (0.143)	0.109 (0.139)	0.195 (0.135)	0.081 (0.131)	0.385*** (0.143)	0.182 (0.118)	0.326*** (0.096)	-0.032 (0.105)	0.004 (0.100)	-835.261
	НОРИТ	0.849*** (0.290)	0.313 (0.283)	0.415 (0.275)	0.420 (0.268)	0.577** (0.291)	0.119 (0.242)	0.546*** (0.194)	-0.087 (0.215)	0.025 (0.203)	-3362.349
	oprobit	0.306** (0.142)	0.228* (0.138)	0.231* (0.133)	0.128 (0.130)	0.255* (0.141)	0.199* (0.117)	0.256*** (0.096)	-0.048 (0.106)	-0.041 (0.098)	-839.096
	НОРИТ	0.561** (0.248)	0.296 (0.242)	0.357 (0.233)	0.280 (0.227)	0.463* (0.247)	0.235 (0.205)	0.321* (0.167)	-0.162 (0.185)	-0.058 (0.172)	-3474.230
Взаимоотношения (n = 519)	oprobit	-0.202 (0.210)	0.072 (0.188)	0.058 (0.183)	-0.116 (0.182)	0.100 (0.198)	0.203 (0.167)	-0.144 (0.132)	-0.066 (0.142)	-0.190 (0.136)	-382.230
	НОРИТ	-0.126 (0.263)	0.091 (0.236)	0.210 (0.231)	0.110 (0.227)	0.129 (0.247)	0.353* (0.209)	-0.238 (0.166)	-0.109 (0.179)	-0.275 (0.171)	-2690.998

Окончание табл. А1

Индикатор здоровья (<i>n</i> = число наблюдений)	Модель	Переменные							Место жительства (1 — город, 0 — село)	Семейный статус (1 — проживает с партнером, 0 — проживает один)	Log-likelihood
		Q1 (беднейшие)	Q2	Q3	Q4	Незаконченное среднее и меньше	Законченное среднее	Возраст			
Конфликты (<i>n</i> = 619)	oprobit	-0.414** (0.194)	-0.112 (0.175)	-0.034 (0.170)	-0.054 (0.165)	0.020 (0.180)	-0.079 (0.146)	0.002 (0.007)	0.014 (0.124)	0.237* (0.139)	-428.927
	NORIT	-0.270 (0.187)	-0.042 (0.170)	-0.191 (0.167)	-0.200 (0.161)	-0.139 (0.176)	-0.116 (0.143)	0.001 (0.007)	-0.073 (0.120)	-0.024 (0.133)	-2672.743
Засыпание (<i>n</i> = 375)	oprobit	0.362** (0.184)	0.125 (0.186)	0.230 (0.174)	-0.091 (0.181)	0.329* (0.190)	0.280* (0.154)	0.024** (0.007)	0.445** (0.131)	0.182 (0.133)	-470.510
	NORIT	0.336 (0.244)	0.152 (0.247)	0.088 (0.243)	-0.188 (0.243)	0.256 (0.251)	0.241 (0.205)	0.021** (0.010)	0.530** (0.176)	0.059 (0.172)	-2147.578
Энергия после пробуждения (<i>n</i> = 410)	oprobit	0.254 (0.182)	0.182 (0.173)	0.207 (0.162)	0.149 (0.166)	0.207 (0.178)	0.062 (0.140)	0.018** (0.007)	0.469** (0.124)	-0.130 (0.127)	-509.345
	NORIT	0.125 (0.252)	0.270 (0.238)	0.151 (0.224)	0.192 (0.230)	0.258 (0.243)	0.092 (0.193)	0.020** (0.010)	0.629** (0.171)	-0.355** (0.174)	-2305.010
Память (<i>n</i> = 691)	oprobit	-0.101 (0.151)	-0.046 (0.140)	0.244* (0.138)	-0.077 (0.133)	0.656** (0.142)	0.313** (0.119)	0.031** (0.006)	0.106 (0.100)	-0.307** (0.110)	-713.513
	NORIT	-0.241 (0.213)	-0.184 (0.197)	0.175 (0.194)	-0.267 (0.187)	0.603** (0.200)	0.311* (0.168)	0.048** (0.009)	0.157 (0.140)	-0.438** (0.154)	-3533.085
Обучение новому (<i>n</i> = 695)	oprobit	-0.034 (0.150)	-0.081 (0.142)	0.264* (0.137)	-0.066 (0.134)	0.710** (0.142)	0.464** (0.119)	0.032** (0.006)	0.119 (0.100)	-0.117 (0.110)	-732.153
	NORIT	-0.097 (0.212)	-0.224 (0.201)	0.275 (0.194)	-0.184 (0.189)	0.807** (0.202)	0.601** (0.170)	0.046** (0.009)	0.193 (0.141)	-0.117 (0.154)	-3482.496
Самообслуживание (<i>n</i> = 514)	oprobit	-0.064 (0.186)	-0.236 (0.190)	0.086 (0.179)	-0.357* (0.188)	0.522** (0.194)	0.375** (0.170)	0.043** (0.008)	0.173 (0.132)	0.167 (0.158)	-414.951
	NORIT	-0.164 (0.370)	-0.623 (0.379)	0.036 (0.356)	-0.729* (0.376)	0.725* (0.387)	0.544 (0.339)	0.093** (0.017)	0.183 (0.262)	-0.046 (0.314)	-2723.564
Внешний вид (<i>n</i> = 570)	oprobit	-0.049 (0.198)	-0.274 (0.207)	0.059 (0.189)	-0.211 (0.192)	0.308 (0.199)	0.278 (0.171)	0.032** (0.008)	0.070 (0.139)	0.031 (0.158)	-338.591
	NORIT	-0.105 (0.377)	-0.665 (0.395)	-0.082 (0.357)	-0.403 (0.365)	0.259 (0.378)	0.424 (0.325)	0.068** (0.017)	0.105 (0.262)	-0.389 (0.299)	-2745.501

Примечание. *, **, *** соответствуют 10, 5 и 1%-ному уровню значимости. В скобках представлены стандартные ошибки. Жирным шрифтом выделены коэффициенты, значимость которых не совпадает для пробит-модели упорядоченного выбора и модели НОРИТ.

Ю. В. Раскина, О. А. Подкорытова, Р. К. Кучаков

Raskina Yu. V., Podkorytova O. A., Kuchakov R. K. Health determinants and the reporting heterogeneity bias in Russia: Anchoring vignettes approach. *Applied Econometrics*, 2022, v. 66, pp. 118–143. DOI: 10.22394/1993-7601-2022-66-118-143

Yulia Raskina

European University at St. Petersburg, St. Petersburg, Russian Federation;
raskina@eu.spb.ru

Olga Podkorytova

European University at St. Petersburg; Saint Petersburg State University, St. Petersburg, Russian Federation;
opodkorytova@eu.spb.ru

Ruslan Kuchakov

European University at St. Petersburg, St. Petersburg, Russian Federation;
rkuchakov@eu.spb.ru

Health determinants and the reporting heterogeneity bias in Russia: Anchoring vignettes approach

We show the reporting heterogeneity bias into all 7 domains of health using the anchoring vignettes approach with the hierarchical ordered probit model. The paper uses the data of the Russian sample of the first wave of the global survey SAGE provided by the World Health Organization Study. The heterogeneity bias correction has changed the significance level of education level, permanent income, rural/urban living, and marital status.

Keywords: determinants of health; reporting heterogeneity bias; anchoring vignettes; HOPIT; Russia; SAGE.

JEL classification: C21; I12.

References

- Gerry C. J., Baydin V. M. (2017). Sources of bias in self-assessed health. *Farmakoekonomika. Sovremennaya farmakoekonomika i farmakoepidemiologiya*, 10 (4), 31–36 (in Russian). DOI: 10.17749/2070-4909.2017.10.4.031-036.
- Kuchakov R. K., Raskina Y. V. (2017). Socioeconomic gradient in health self-assessments of the elderly population in Russia and China: Results of hierarchical ordinal probit model. *Finansy i biznes*, 13 (4) 22–46 (in Russian).
- Kaneva M., Baidin V. (2018). Heterogeneity in reporting self-assessed health of the Russians. *Applied Econometrics*, 3 (51), 102–125 (in Russian).
- Paul P., Valtonen H., Kovtun N. V. (2019). Socioeconomic inequalities in health in the post-Soviet Russia. *Population*, 22 (1), 61–78 (in Russian). DOI: 10.19181/1561-7785-2019-00005.
- Khabibullina A. R., Aleksandrova E. A., Aistov A. V. (2012). Estimates of health inequality among Russians. *Voprosy Ekonomiki*, 3, 117–138 (in Russian). DOI: 10.32609/0042-8736-2021-3-117-138.
- Angelini V., Cavapozzi D., Corazzini L., Paccagnella O. (2014). Do Danes and Italians rate life satisfaction in the same way? Using vignettes to correct for individual-specific scale biases. *Oxford Bulletin of Economics and Statistics*, 76 (5), 643–666. DOI: 10.1111/obes.12039.

Bago d'Uva T., Doorslaer van E., Lindeboom M., O'Donnell O. (2008). Does reporting heterogeneity bias the measurement of health disparities? *Health Economics*, 17 (3), 351–375. DOI: 10.1002/hec.1269.

Bago d'Uva T., Lindeboom M., O'Donnell O., Doorslaer van E. (2011). Slipping anchor? Testing the vignettes approach to identification and correction of reporting heterogeneity. *Journal of Human Resources*, 46 (4), 875–906. DOI: 10.1353/jhr.2011.0005.

Baker M., Stabile M., Deri C. (2004). What do self-reported, objective, measures of health measure? *Journal of Human Resources*, 39 (4), 1067–1093. DOI: 10.2307/3559039.

Bambra C., Popham F. (2010). Worklessness and regional differences in the social gradient in general health: Evidence from the 2001 English census. *Health & Place*, 16 (5), 1014–1021. DOI: 10.1016/j.healthplace.2010.06.006.

Burström B. (2002). Increasing inequalities in health care utilisation across income groups in Sweden during the 1990s? *Health Policy*, 62 (2), 117–129. DOI: 10.1016/s0168-8510(02)00016-7.

Cutler D. M., Lleras-Muney A. (2006). Education and health: Evaluating theories and evidence. *Working Paper* 12352. National Bureau of Economic Research. DOI: 10.3386/w12352.

DeSalvo K. B., Bloser N., Reynolds K., He J., Muntner P. (2006). Mortality prediction with a single general self-rated health question. *Journal of General Internal Medicine*, 21 (3), 267–275. DOI: 10.1111/j.1525-1497.2005.00291.x.

Doorslaer van E., Koolman X. (2004). Explaining the differences in income-related health inequalities across European countries. *Health Economics*, 13 (7), 609–628. DOI: 10.1002/hec.918.

Eikemo T. A., Bambra C., Joyce K., Dahl E. (2008). Welfare state regimes and income-related health inequalities: A comparison of 23 European countries. *European Journal of Public Health*, 18 (6), 593–599. DOI: 10.1093/eurpub/ckn092.

Ferguson B. D., Tandon A., Gakidou E., Murray C. J. (2003). Estimating permanent income using indicator variables. In: C. J. L. Murray, D. B. Evans (eds.). *Health Systems Performance Assessment: Debates, Methods and Empiricism*, 747–760. Geneva: World Health Organization.

Grol-Prokopczyk H. (2014). Age and sex effects in anchoring vignette studies: Methodological and empirical contributions. *Survey Research Methods*, 8 (1), 1–17. DOI: 10.18148/srm/2014.v8i1.5133.

Gorshkov I., Zharinova A. (2016). The effect of education on health perception and health related behavior: Evidence from the Russian Federation. DOI: 10.2139/ssrn.2827370.

Groot W., Den Brink van H. M. (2007). The health effects of education. *Economics of Education Review*, 26 (2), 186–200. DOI: 10.1016/j.econedurev.2005.09.002.

Gupta N. D., Jürges H. (2012). Do workers underreport morbidity? The accuracy of self-reports of chronic conditions. *Social Science & Medicine*, 75 (9), 1589–1594. DOI: 10.1016/j.socscimed.2012.07.013.

Gupta N. D., Kristensen N., Pozzoli D. (2010). External validation of the use of vignettes in cross-country health studies. *Economic Modelling*, 27 (4), 854–865. DOI: 10.1016/j.econmod.2009.11.007.

Hopkins D. J., King G. (2010). Improving anchoring vignettes: Designing surveys to correct interpersonal incomparability. *Public Opinion Quarterly*, 74 (2), 201–222. DOI: 10.1093/poq/nfq011.

Idler E. L., Benyamini Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior*, 38 (1), 21–37. DOI: 10.2307/2955359.

Jones A. M., Rice N., Bago d'Uva T., Balia S. (2012). *Applied health economics*. Routledge. DOI: 10.4324/9780203102411.

Jürges H. (2007). True health vs response styles: Exploring cross-country differences in self-reported health. *Health Economics*, 16 (2), 163–178. DOI: 10.1002/hec.1134.

Jylhä M. (2009). What is self-rated health and why does it predict mortality? Towards a unified conceptual model. *Social Science & Medicine*, 69 (3), 307–316. DOI: 10.1016/j.socscimed.2009.05.013.

Jylhä M., Volpato S., Guralnik J. M. (2006). Self-rated health showed a graded association with frequently used biomarkers in a large population sample. *Journal of Clinical Epidemiology*, 59 (5), 465–471. DOI: 10.1016/j.jclinepi.2005.12.004.

Kapteyn A., Smith J. P., van Soest A. H. O. (2007). Vignettes and self-reports of work disability in the United States and the Netherlands. *American Economic Review*, 97 (1), 461–473. DOI: 10.1257/aer.97.1.461.

Kerkhofs M., Lindeboom M. (1995). Subjective health measures and state dependent reporting errors. *Health Economics*, 4 (3), 221–235. DOI: 10.1002/hec.4730040307.

King G., Murray C. J., Salomon J. A., Tandon A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review*, 98 (1), 191–207. DOI: 10.1017/S000305540400108X.

King G., Wand J. (2007). Comparing incomparable survey responses: Evaluating and selecting anchoring vignettes. *Political Analysis*, 15 (1), 46–66. DOI: 10.1093/pan/mpl011.

Krieger N., Chen J. T., Waterman P. D., Rehkopf D. H., Subramanian S. V. (2003). Race/ethnicity, gender, and monitoring socioeconomic gradients in health: A comparison of area-based socioeconomic measures — the public health disparities geocoding project. *American Journal of Public Health*, 93 (10), 1655–1671. DOI: 10.2105/ajph.93.10.1655.

Lindeboom M., Doorslaer van E. (2004). Cut-point shift and index shift in self-reported health. *Journal of Health Economics*, 23 (6), 1083–1099. DOI: 10.1016/j.jhealeco.2004.01.002.

Lokshin M., Ravallion M. (2008). Testing for an economic gradient in health status using subjective data. *Health Economics*, 17 (11), 1237–1259. DOI: 10.1093/ije/30.3.600.

Manor O., Matthews S., Power C. (2001). Self-rated health and limiting longstanding illness: Interrelationships with morbidity in early adulthood. *International Journal of Epidemiology*, 30 (3), 600–607. DOI: 10.1093/ije/30.3.600.

Marmot M., Friel S., Bell R., Houweling T. A., Taylor S. (2008). Closing the gap in a generation: Health equity through action on the social determinants of health. *The Lancet*, 372 (9650), 1661–1669. DOI: 10.1016/S0140-6736(08)61690-6.

Martinson M. L. (2012). Income inequality in health at all ages: A comparison of the United States and England. *American Journal of Public Health*, 102 (11), 2049–2056. DOI: 10.2105/AJPH.2012.300929.

Miao J., Wu X. (2016). Urbanization, socioeconomic status and health disparity in China. *Health & Place*, 42, 87–95. DOI: 10.1016/j.healthplace.2016.09.008.

Molina T. (2017). Adjusting for heterogeneous response thresholds in cross-country comparisons of self-reported health. *The Journal of the Economics of Ageing*, 10, 1–20. DOI: 10.1016/j.jeoa.2017.08.002.

Montez J. K., Berkman L. F. (2013). Trends in the educational gradient of mortality among US adults aged 45 to 84 years: Bringing regional context into the explanation. *American Journal of Public Health*, 104 (1), 82–90. DOI: 10.2105/AJPH.2013.301526.

Murray C. J., Ozaltin E., Tandon A., Salomon J., Sadana R., Chatterji S. (2003). Empirical evaluation of the anchoring vignette approach in health surveys. In: C. J. L. Murray, D. B. Evans (eds.). *Health Systems Performance Assessment: Debates, Methods and Empiricism*, 369–399. Geneva: World Health Organization.

Newacheck P. W., Hung Y. Y., Jane Park M., Brindis C. D., Irwin Jr C. E. (2003). Disparities in adolescent health and health care: Does socioeconomic status matter? *Health Services Research*, 38 (5), 1235–1252. DOI: 10.1111/1475-6773.00174.

Nielsen A. B. S., Siersma V., Conradsen Hiort L., Drivsholm T., Kreiner S., Hollnagel H. (2008). Self-rated general health among 40-year-old Danes and its association with all-cause mortality at 10-, 20-, and 29 years follow-up. *Scandinavian Journal of Public Health*, 36 (1), 3–11. DOI: 10.1177/1403494807085242.

Oksuzyan A., Daňko M. J., Caputo J., Jasilionis D., Shkolnikov V. M. (2019). Is the story about sensitive women and stoical men true? Gender differences in health after adjustment for reporting behavior. *Social Science & Medicine*, 228, 41–50. DOI: 10.1016/j.socscimed.2019.03.002.

Paul P., Valtonen H. (2016). Inequalities in perceived health in the Russian Federation. 1994–2012. *BMC Public Health*, 16 (1), 1–12. DOI: 10.1186/s12889-016-2810-x.

Peracchi F., Rossetti C. (2013). The heterogeneous thresholds ordered response model: Identification and inference. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176 (3), 703–722. DOI: 10.1111/j.1467-985X.2012.01066.x.

Rice N., Robone S., Smith P. (2011). Analysis of the validity of the vignette approach to correct for heterogeneity in reporting health system responsiveness. *The European Journal of Health Economics*, 12 (2), 141–162. DOI: 10.1007/s10198-010-0235-5.

Riva M., Curtis S., Gauvin L., Fagg J. (2009). Unravelling the extent of inequalities in health across urban and rural areas: Evidence from a national sample in England. *Social Science & Medicine*, 68 (4), 654–663. DOI: 10.1016/j.socscimed.2008.11.024.

Rossouw L., Bago d’Uva T., Doorslaer van E. (2018). Poor health reporting? Using anchoring vignettes to uncover health disparities by wealth and race. *Demography*, 55 (5), 1935–1956. DOI: 10.1007/s13524-018-0709-0.

Van Soest A. H. O., Delaney, L., Harmon C., Kapteyn A., Smith J. P. (2011). Validating the use of anchoring vignettes for the correction of response scale differences in subjective questions. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174 (3), 575–595. DOI: 10.1111/j.1467-985X.2011.00694.x.

Van Soest A. H. O., Vonkova H. (2014). Testing the specification of parametric models by using anchoring vignettes. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 177 (1), 115–133. DOI: 10.1111/j.1467-985X.2012.12000.x.

Stafford M., Cummins S., Macintyre S., Ellaway A., Marmot M. (2005). Gender differences in the associations between health and neighbourhood environment. *Social Science & Medicine*, 60 (8), 1681–1692. DOI: 10.1016/j.socscimed.2004.08.028.

Terraneo M. (2015). Inequities in health care utilization by people aged 50+: Evidence from 12 European countries. *Social Science & Medicine*, 126, 154–163. DOI: 10.1016/j.socscimed.2014.12.028.

Valentine N., Verdes-Tennant E., Bonsel G. (2015). Health systems’ responsiveness and reporting behaviour: Multilevel analysis of the influence of individual-level factors in 64 countries. *Social Science & Medicine*, 138, 152–160. DOI: 10.1016/j.socscimed.2015.04.022.

Scott N. W. et al. (2010). Differential item functioning (DIF) analyses of health-related quality of life instruments using logistic regression. *Health and Quality of Life Outcomes*, 8 (1), 1–9. DOI: 10.1186/1477-7525-8-81.

Schneider U., Pfarr C., Schneider B. S., Ulrich V. I. (2012). I feel good! Gender differences and reporting heterogeneity in self-assessed health. *The European Journal of Health Economics*, 13 (3), 251–265. DOI: 10.1007/s10198-011-0301-7.

World Health Organization. (2006). WHO SAGE survey manual: The WHO study on global ageing and adult health (SAGE). Geneva: World Health Organization.

Received 07.09.2021; accepted 17.05.2022.