

Прикладная эконометрика, 2020, т. 60, с. 115–138.

Applied Econometrics, 2020, v. 60, pp. 115–138.

DOI: 10.22394/1993-7601-2020-60-115-138

К. Н. Макшанчиков¹

Расходы россиян на спорт: эконометрический анализ по данным опроса Левада-Центра

Рассматривается принятие взрослыми индивидами решений об оплате занятий спортом и о затратах на непрофессиональные занятия спортом в оздоровительных целях. Для эмпирической оценки используются данные проведенного в 2017 г. опроса аналитического центра Юрия Левады об отношении людей к собственному здоровью и о качестве медицинской помощи в России. При оценивании применяются вероятностные модели бинарного выбора, модели Хекмана, двухшаговые модели Хекмана, а также модели Ньюи с полупараметрическим способом коррекции смещения отбора, позволяющие ослабить предположение о нормальном распределении случайных ошибок. Результаты исследования свидетельствуют о значимой взаимосвязи расходов на спорт с доходами индивидов, различающейся для разных половозрастных групп.

Ключевые слова: занятия спортом; затраты на спорт; инвестиции в человеческий капитал; образ жизни; непрофессиональный спорт.

JEL classification: C21; C31; I12; Z20.

1. Введение

Одним из важнейших аспектов здорового образа жизни является физическая активность, включающая в том числе занятия спортом. Недостаточный уровень физической активности и, как следствие, болезни, связанные с избыточным весом, принадлежат к группе основных причин смертности взрослого населения (WHO, 2010). Спорт, как один из видов физической активности, предоставляет индивиду возможность укреплять свое здоровье, работодателю — сокращать проблемы, возникающие из-за пропусков по болезни и сниженной продуктивности, а государству — обеспечивать развитие человеческого потенциала и улучшать качество жизни граждан (Cawley, 2004; Miles, 2007; Downward, Rasciute, 2011). Кроме того, занятия людей спортом, помимо улучшения состояния здоровья, сопровождаются множеством положительных эффектов, таких как позитивное влияние на достижения в области образования (Pfeifer, Cornelißen, 2010), на профессиональном поприще (Lechner, 2009), повышение уровня социализации (Downward, Riordan, 2007), падение преступности (Caguso, 2011).

¹ Макшанчиков Константин Николаевич — Национальный исследовательский университет «Высшая школа экономики», Москва; kmakshanchikov@hse.ru.

Последние 15 лет отмечены положительными изменениями в отношении российских граждан к спорту: доля регулярно занимающихся спортом людей увеличилась почти втрое и приблизилась к 30%, однако все еще остается ниже уровня 40–60%, характерного для развитых стран (Шишкин и др., 2017). Тем не менее, исследования факторов, связанных с размером расходов на спорт, на российских данных обнаружить не удалось, что и послужило мотивом для настоящей статьи. Цель данной работы — устранить пробел в исследовании принятия взрослыми индивидами решений об оплате занятий спортом и физкультурой и определить взаимосвязь этих решений с социально-демографическими и экономическими характеристиками индивидов. Отдельное внимание в работе уделяется выбору наилучшего подхода для оценки размера оплаты занятий спортом на российских данных.

Спортивная активность может осуществляться в рамках профессиональной деятельности (если индивид является спортсменом), в качестве способа проведения досуга и в оздоровительных целях. Во всех перечисленных случаях спортивная активность может (хоть и не всегда) способствовать укреплению здоровья. Тем не менее, в данной работе будет изучаться преимущественно спортивная активность индивидов, изначально направленная на укрепление здоровья. Этим объясняется и дальнейший выбор теоретических концепций, и база данных для эмпирического анализа.

Работа состоит из пяти разделов. Второй раздел посвящен обзору теоретических работ в области оплаты занятий спортом на индивидуальном уровне, а также анализу экономических факторов, влияющих на вероятность их оплаты. В третьем и четвертом разделах содержится эмпирическая часть исследования: описание объекта исследования, гипотезы относительно ожидаемых результатов, результаты проделанной работы. Завершается статья заключением с основными выводами.

2. Обзор современного состояния изучаемой проблемы

Первые эмпирические исследования, посвященные досугу и изучению занятий людей спортом, были проведены в США. Для исследования решений индивидов заниматься спортом и периодичности этих занятий использовалась двухэтапная эконометрическая модель (Cicchetti et al., 1969). На европейском уровне подобные исследования показали значительное сходство в характере участия в спорте в разных европейских странах (Downward et al., 2011).

С тех пор модели решения об участии в спорте усложнились: на данный момент существуют исследования, использующие логистические и двухшаговые модели Хекмана, а также многофакторный анализ (Breuer, Wicker, 2008; Downward, 2007; Downward, Riordan, 2007; Farrell, Shields, 2002; Humphreys, Ruseski, 2006, 2007). Имеющиеся исследования в качестве предмета анализа рассматривают: вероятность участия (в том числе в разных видах спорта); интенсивность занятий (как часто люди занимаются спортом и сколько времени тратят на каждое занятие, или с какой нагрузкой занимаются — слабой, умеренной, интенсивной); расходы на спорт (могут быть индивидуальными и семейными). Предметом анализа данной работы являются расходы на спорт на индивидуальном уровне.

Несмотря на различие подходов, можно обобщить результаты о роли экономических, индивидуальных и социальных переменных в принятии решений о занятии спортом. В большинстве эмпирических исследований вероятность участия в спорте уменьшается с возрастом (Barber, Navitz, 2001; Breuer, Wicker, 2009; Downward, 2007; Downward, Riordan, 2007;

Farrell, Shields, 2002; Humphreys, Ruseski, 2006; Moens, Scheerder, 2004; Scheerder et al., 2005; Wicker et al., 2009). В то же время на данных по Италии выявлено, что максимальная вероятность участия в спорте достигается в 33 года (García et al., 2011). Это означает, что кривая зависимости вероятности занятия физической активностью от возраста имеет вид перевернутой U-образной кривой с максимумом в точке 33 года.

Степень вовлеченности в спорт может быть связана с биологическими и физическими ограничениями и переходом к видам спортивной деятельности, предпочитаемых старшими возрастными группами (Barber, Havitz, 2001). Вероятность занятия большинством видов спорта падает с возрастом, за редким исключением йоги и ходьбы — ходьба является положительно связанной с возрастом (Strawinski, 2010; Humphreys, Ruseski, 2007; Lera-López, Rapún-Gárate, 2007). По некоторым исследованиям, интенсивность участия в спорте увеличивается с возрастом (Humphreys, Ruseski, 2006; Lera-López, Rapún-Gárate, 2007).

Существует устоявшееся мнение о том, что мужчины, в целом, не только участвуют в спорте больше, чем женщины (Downward, 2007; Humphreys, Ruseski, 2006, 2007; Lera-López, Rapún-Gárate, 2007; Moens, Scheerder, 2004; Wilson, 2002), но также показывают более высокую интенсивность занятий (Barber, Havitz, 2001; Humphreys, Ruseski, 2006, 2007). Гендерные различия могут объясняться различными факторами, например, биологическими, культурными и социальными, отражающими различия в семейных обязанностях, а также неодинаковостью в социальных ожиданиях и профессиональном статусе. Таким образом, важность гендерной проблематики при разделении труда домашних хозяйств в рамках модели распределения времени и тот факт, что спорт неразрывно связан с гендерной идентичностью и ролями в обществе, побуждают проводить анализ для мужчин и для женщин по отдельности (Downward et al., 2014).

Традиционно подчеркивается значение экономических переменных при исследовании участия в спортивных мероприятиях. Чаще всего для оценки экономического положения индивида используются переменные, отражающие его доход. В литературе приводятся доказательства того, что более низкий доход может действовать как барьер для занятий спортом (Breuer, Wicker, 2009; Farrell, Shields, 2002; Humphreys, Ruseski, 2006, 2007; Lera-López, Rapún-Gárate, 2007; Stempel, 2005; Wicker et al., 2009; Wilson, 2002), а увеличение доходов способствует росту вероятности занятий физической активностью (Humphreys, Ruseski, 2006; Strawinski, 2010; Downward, Rasciute, 2010). Тем не менее, иногда доход не имеет никакого влияния на интенсивность участия в спортивных мероприятиях или имеет отрицательное влияние (Downward, Riordan, 2007; Humphreys, Ruseski, 2006). По данным опроса Левада-Центра 2011 г. была обнаружена устойчивая положительная связь между доходами и вероятностью занятий спортом (Засимова, Локтев, 2016).

Еще одним важным фактором в экономическом анализе спроса на спорт является время. Поскольку временные ресурсы ограничены, любое увеличение времени на занятия спортом всегда будет конкурировать с временем на отдых, работу и другие виды деятельности. Сдерживающие временные факторы воспринимаются как один из наиболее значимых барьеров для физической активности и участия в спортивных мероприятиях (Alexandris, Cartoll, 1999). Влияние переменной времени обычно анализируется через три прокси-переменные: профессию и профессиональный статус, размер семьи, семейное положение. Меньшее участие в спорте в целом было найдено среди определенных профессиональных слоев в низших социально-экономических группах (Lera-López, Rapún-Gárate, 2007). С ростом времени, проведенного на работе, снижалась вероятность занятий спортом (Breuer, Wicker, 2008; Downward, 2007).

Исследования также показывают, что характеристики домохозяйств могут быть важны в объяснении участия в спортивных мероприятиях. Размер домохозяйства отрицательно связан с занятиями спортом (Downward, 2007; Humphreys, Ruseski, 2006, 2007; Scheerder et al., 2005). В то же время отмечается, что есть зависимость от рассматриваемого вида спорта. Наличие детей может ограничить время, доступное взрослым для занятий такими видами спорта, как аэробика и бег, и увеличить их участие в ориентированных на ребенка видах спорта, среди которых футбол или плавание (Downward, 2004). Время на уход за детьми или родственниками негативно влияет на регулярные занятия спортом (Breuer et al., 2010), но некоторые исследования отмечают, что дети способствуют росту вероятности участия родителей в спорте (Strawinski, 2010). Женатые люди меньше участвуют в спортивных и физкультурных мероприятиях (Humphreys, Ruseski, 2006), хотя имеются различия по видам деятельности и интенсивности спортивных мероприятий (Humphreys, Ruseski, 2006, 2007).

Более высокий уровень образования может привести к повышению осведомленности о пользе и значении спорта, а также может быть связан с более высокой заработной платой. Образование также связывают с формированием привычки к занятиям спортом во время обучения в вузе, поскольку у студентов имеется простой и недорогой доступ к спортивной инфраструктуре. Таким образом, можно ожидать положительную связь между образованием и участием в спортивных мероприятиях, что подтверждается в эмпирических исследованиях (Downward, 2007; Humphreys, Ruseski, 2006, 2007; Stempel, 2005; Wilson, 2002). Однако частота занятий спортом чаще всего отрицательно связана с уровнем образования (Downward, Riordan, 2007; Humphreys, Ruseski, 2006).

Некоторые исследования рассматривают влияние самооценки здоровья в качестве важного фактора участия в спортивных мероприятиях. Эта переменная, вероятно, должна иметь положительное влияние на частоту занятий спортом (Downward, Riordan, 2007; Humphreys, Ruseski, 2006, 2007). Однако для некоторых видов спорта, таких как плавание, езда на велосипеде и футбол, этот фактор может быть незначимым (Downward, 2007). Иногда плохое здоровье может подвигнуть индивида к занятиям определенными видами спорта по совету врача (Farrell, Shields, 2002).

Наконец, многие исследования включают фактор доступности спортивных объектов. В самом деле, меньший доступ к определенным видам спортивных сооружений в сельской местности, чем в пригороде или городах, негативно влияет на занятие людей спортом (Andreff, 2001), что может оказать отрицательное влияние на общий уровень участия в спортивных мероприятиях (Moens, Scheerder, 2004; Scheerder et al., 2005). На российских данных была обнаружена позитивная связь наличия спортивной инфраструктуры по месту работы или учебы и вероятности занятий людей спортом (Засимова, Локтев, 2016).

Решение об оплате занятий спортом, вероятно, связано с решением о занятии спортом. Поэтому в исследованиях рассматриваются двухшаговые модели, в которых индивиды на первом этапе принимают решение об оплате занятий спортом, а на втором — о размере оплаты занятий спортом.

В исследованиях предпринимаются попытки объяснить различия в расходах на спорт в зависимости от таких характеристик домашних хозяйств, как место проживания, пол, возраст, образование, рабочий и социальный статус, размер домохозяйства, число детей (Pawlowski, Breuer, 2011). Исследователи отмечают, что возраст не оказывает значимого влияния на общие спортивные расходы (Pawlowski, Breuer, 2011, 2012; Thibaut et al., 2014; Wicker et al., 2010), однако возраст становится значимым показателем при исследовании определенных видов

расходов, например, на спортивную одежду (Scheerder et al., 2011). Пол индивида по-разному влияет на затраты на спорт: у женщин рост затрат связан с членством в спортивных клубах (Wicker et al., 2010), а мужчины тратят больше денег на спортивную одежду и снаряжение (Scheerder et al., 2011). Установлено, что семейное положение не связано с затратами на спорт (Pawlowski, Breuer, 2011, 2012), а рабочий и социальный статус оказывает существенное влияние: «белые воротнички» тратят больше на спортивные товары, а безработные тратят в целом на спорт меньше, чем «синие воротнички» (Pawlowski, Breuer, 2012).

В ряде исследований выявлено положительное влияние размера домохозяйства на спортивные расходы (Pawlowski, Breuer, 2011, 2012), в то время как в других его влияние оказалось незначимым (Scheerder et al., 2011; Thibaut et al., 2014). Возраст детей определяет размеры затрат на спорт: присутствие маленьких детей (до 6 лет) приводит к снижению расходов семьи на спорт (Thibaut et al., 2014), а наличие детей старшего возраста (от 6 до 18 лет) — к их увеличению (Pawlowski, Breuer, 2011, 2012). Место проживания индивидов существенно влияет на размер спортивных расходов домохозяйств: они имеют тенденцию к росту с увеличением размера населенного пункта (Pawlowski, Breuer, 2011, 2012). Доходы домашних хозяйств включаются в большинство исследований, и их влияние признается значимо положительным.

Проведенный анализ позволяет предположить, что многие из факторов, оказавшихся значимыми в предыдущих исследованиях, могут оказаться значимыми и для России.

3. Данные

Эмпирический анализ осуществляется на базе данных опроса аналитического центра Юрия Левады об отношении людей к собственному здоровью и о качестве медицинской помощи в России, проведенного в 2017 г. Выборка репрезентативна на национальном уровне и содержит данные по 4006 индивидам в возрасте от 16 лет.

На вопрос о занятиях спортом и физкультурой в оздоровительных целях около двух третей опрошенных (2705 респондентов или 67.5% выборки) сообщили, что практически не занимаются спортом в оздоровительных целях. Среди них 61 респондент (1.5%) отметили, что занимаются спортом только профессионально, поэтому они в данном исследовании не учитывались. Всего занимаются спортом в оздоровительных целях 1229 человек (30.7%), что на 4.2% больше, чем в аналогичном опросе 2011 г. (Засимова, Локтев 2016). Далее в работе под термином «занятия спортом» будут пониматься именно занятия спортом и физкультурой в оздоровительных целях. На основе анализа предыдущих исследований были сформированы следующие гипотезы о влиянии факторов на вероятность оплаты и размер оплаты занятий спортом:

- H1: *вероятность оплаты занятий спортом положительно связана с доходами индивида;*
H2: *размер оплаты занятий спортом положительно связан с доходами индивида.*

Для проверки гипотез построены три зависимые переменные, отвечающие за занятия спортом, оплату и размер оплаты занятий спортом:

- *sport* — бинарная переменная, равная 1, если индивид занимается спортом в оздоровительных целях;
- *sportpay* — бинарная переменная, равная 1, если индивид тратит деньги на занятия спортом в оздоровительных целях;

- *lnsps* — переменная, равная нулю, если индивид не тратит деньги на занятия спортом, и логарифму величины расходов на занятия спортом иначе.

В опросе содержался вопрос о том, какую сумму (в рублях) тратит индивид ежемесячно на занятия спортом, не учитывая расходы на спортивные товары и одежду. На вопрос о размере оплаты занятий спортом положительно ответили 309 человек, что составляет четвертую часть от занимающихся спортом в оздоровительных целях.

Обзор литературы позволил выделить факторы, связанные с решением индивида об оплате занятий спортом и ее размере. Ниже кратко охарактеризуем основные переменные и подробнее остановимся на тех из них, построение которых требует дополнительного объяснения.

Доход. В базе содержится несколько вопросов, касающихся доходов индивидов. Первый тип вопросов требует от индивида конкретных значений дохода на основной работе, премий, дохода за дополнительную работу. Однако людям свойственно занижать свой реальный заработок, уходить от ответа. На прямой вопрос о размере заработка на основной работе ответили 1504 индивида, на дополнительной работе — 181 индивид. Используя эти данные, оценить доход индивида удалось лишь для 37.5% выборки. Второй тип вопросов касался дохода домохозяйства. На прямой вопрос о размере доходов домохозяйства ответили 2183 индивида. В случае отказа от ответа индивиду предлагалось оценить размер доходов домохозяйства с помощью категорий. Агрегируя ответы на вопросы о суммарном размере денежных доходов домохозяйств и учитывая размер домохозяйств, удалось оценить среднедушевой доход 3375 индивидов (84% выборки). Третий тип вопросов — категориальный. На него не ответили всего 10 индивидов. Вопрос распределяет индивидов по 6 категориям, от 1 (самые бедные) до 6 (самые богатые). В связи с незначительным размером категорий самых бедных и самых богатых, было решено объединить их с соседними категориями. В моделях используются следующие переменные:

lnincome — логарифм дохода индивида;

lnahhinc — логарифм дохода на одного члена домохозяйства;

incotegroup — категория доходов индивида.

Здоровье. Принятие индивидом решения о занятиях спортом может зависеть от его оценки собственного здоровья (Downward et al., 2007; Humphreys, Ruseski, 2006, 2007; Засимова, Локтев, 2016). Ожидается положительное влияние этого фактора на занятия спортом. На основе ответов респондентов на вопросы о самооценке здоровья и наличии/отсутствии инвалидности были построены две бинарные переменные, отвечающие за хорошее и плохое здоровье индивида:

ghealth — бинарная переменная, равная 1, если у индивидов не было инвалидности и они оценивали свое здоровье как хорошее или очень хорошее;

bhealth — бинарная переменная, равная 1, если у индивидов имелась инвалидность или они отмечали, что считают свое здоровье плохим или очень плохим.

Вредные привычки. В работе проверяется наличие связи между курением и потреблением алкоголя и вероятностью принятия решения об оплате и размере оплаты спортивных занятий. Для этого созданы две бинарные переменные, отвечающие за регулярное потребление:

smoke — бинарная переменная, равная 1, если индивид курит как минимум 1 раз в день;

drink — равна 1, если индивид употребляет алкогольные напитки чаще, чем 2 раза в неделю.

Окружение. При оплате занятий спортом довольно часто предоставляются групповые (семейные) абонементы, а также скидки для работодателей. Поэтому логично предположить, что если индивид занимается спортом вместе со своими коллегами или родственниками, то он тратит меньше, чем при самостоятельных занятиях. Ожидалось, что если вместе с индивидом занимается спортом кто-то из близких или друзей, то он с большей вероятностью склонен платить за занятия спортом, нежели тот, кто занимается в одиночку.

Кроме того, в модели входят переменные, характеризующие пол, возраст индивида, его профессиональную категорию, размер домохозяйства, наличие детей, семейный статус, уровень образования, наличие рядом с местом проживания, учебы или работы спортивной инфраструктуры, тип населенного пункта и федерального округа, в котором он проживает. Описательные статистики непрерывных и дискретных независимых переменных представлены в Приложении в таблицах П1 и П2 соответственно.

4. Эконометрическое моделирование

Построение эконометрических моделей проводится на выборках, включающих людей, занимающихся спортом в оздоровительных целях и не включающих тех, для кого занятия спортом являются частью их профессии. Для каждого типа моделей рассматриваются три спецификации, отличающиеся друг от друга включением переменной дохода:

- (1) с переменной $\ln income$ — логарифм дохода индивида;
- (2) с переменной $\ln hhinc$ — логарифм дохода на одного члена домохозяйства;
- (3) с переменной $incomegroup$ — категория доходов индивида.

Исследование зависимости при помощи разных переменных было необходимо ввиду наличия проблемы с пропусками по первым двум переменным (и, следовательно, снижением размера выборки и смещением в сторону низкодоходных групп из-за отказа от ответа или занижения своего дохода индивидами с высокими доходами), а также для проверки устойчивости результатов.

Кроме того, могут иметь место различия во влиянии переменных для мужчин и женщин. Поэтому после оценки моделей на общей выборке дополнительно проверяется гипотеза о возможности оценивания совместной модели с фиктивной переменной гендерности.

На первом этапе для оценивания вероятности оплаты занятий спортом используется пробит-модель с учетом смещения выборки, поскольку индивиды, которые не занимаются спортом в оздоровительных целях, не тратят на это деньги. Главное предположение модели состоит в том, что и уравнение отбора (решение о занятиях спортом в оздоровительных целях), и основное уравнение (решение об оплате занятий спортом) являются бинарными. Согласно предположениям модели, зависимая переменная (решение об оплате занятий спортом) наблюдаема, если:

$$y_i^{probit} = I(y_i^* > 0), \quad (1)$$

где $y_i^* = X_i'\beta + \varepsilon_{i1}$ — ненаблюдаемая вероятность оплаты занятий спортом, и выполняется условие отбора следующего вида:

$$y_i^{select} = I(X_i'\beta + \varepsilon_{i2} > 0), \quad (2)$$

где $\varepsilon_{i1} \sim N(0; 1)$, $\varepsilon_{i2} \sim N(0; 1)$, $\text{corr}(\varepsilon_{i1}, \varepsilon_{i2}) = \rho$; X_i — контрольные переменные; y_i^{select} — индивид занимается спортом в оздоровительных целях; β — вектор коэффициентов, характеризующий влияние объясняющих переменных; ε_{i1} , ε_{i2} — случайные ошибки; $i = 1, \dots, N$ — число наблюдений (индивидов), $I(\cdot)$ — индикаторная функция, принимающая значение 1, если условие в скобках выполнено, и 0 в противном случае. Если $\rho = 0$, то индивиды отобраны случайным образом, и достоверные данные могут быть получены с помощью простой пробит-модели.

Далее оценивается размер оплаты занятий непрофессиональным спортом в оздоровительных целях с помощью модели Хекмана, которая позволяет предположить наличие двух взаимосвязанных решений. В качестве уравнения отбора принимается решение индивида о занятиях спортом в оздоровительных целях, а в качестве основного уравнения — размер оплаты занятий спортом. ненаблюдаемые вероятность оплаты занятий спортом и их размер задаются уравнениями:

$$y_{ji}^* = X'_{ji} \beta_j + \varepsilon_{ji}, \quad j = 1, 2, \quad (3)$$

$$y_{2i} = \begin{cases} y_{2i}^*, & \text{если } y_{1i}^* > 0, \\ 0, & \text{если } y_{1i}^* \leq 0, \end{cases} \quad (4)$$

где y_{1i} — бинарная переменная на занятия спортом, принимающая значение 1 для занимающихся спортом и 0 для не занимающихся, а y_{2i} — наблюдаемые расходы на спорт. Используемая модель Хекмана оценивается методом максимального правдоподобия и с помощью двухшаговой процедуры (Heckman, 1976). Для реализации последней необходимо на первом шаге оценить уравнение участия с помощью пробит-модели, используя полученные оценки $\hat{\beta}_1$, рассчитать обратное отношение Миллса $\varphi(X'_{1i} \hat{\beta}_1) / \Phi(X'_{1i} \hat{\beta}_1)$, где φ и Φ — функции плотности и распределения стандартного нормального закона соответственно, а затем на втором шаге оценить основное уравнение методом наименьших квадратов:

$$y_i = X'_{2i} \beta_2 + \rho \sigma \hat{\lambda}_i + \varepsilon_i, \quad (5)$$

где $\hat{\lambda}_i$ — оценка обратного отношения Миллса, ε_i — случайная ошибка. Одним из предположений модели является совместное нормальное распределение случайных ошибок:

$$\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{bmatrix} \right), \quad (6)$$

в рамках которого математическое ожидание зависимой переменной имеет вид:

$$E(y_{2i}) = E(y_{2i}^* | y_{1i} = 1) = E(y_{2i}^* | y_{1i}^* \geq 0) = X'_{2i} \beta_2 + \rho \sigma \lambda_i. \quad (7)$$

Однако среди значений основной переменной (расходов на спорт) много нулей — это расходы индивидов, которые занимаются спортом бесплатно. Поэтому предположение о совместном нормальном распределении случайных ошибок в уравнении участия и в основном уравнении является необоснованным. Для решения этой проблемы рассматриваются модели, отчасти корректирующие смещение отбора — в качестве уравнения участия выбирается факт оплаты занятий спортом. Результаты моделирования показывают, что неслучайный отбор имеет место не среди тех, кто платит или нет, а среди тех, кто занимается

или не занимается спортом. Поэтому оценивать уравнение размера оплаты непрофессиональных занятий спортом в оздоровительных целях можно с помощью метода наименьших квадратов.

Незначимость связи между случайными ошибками, полученная при построении моделей методом Хекмана, может быть обусловлена некорректным результатом теста о независимости случайных ошибок вследствие ошибки в спецификации их совместного распределения. Чтобы ослабить допущение о совместном нормальном распределении случайных ошибок, используется полупараметрический двухшаговый подход Ньюи (Newey, 2009; De Luca, Petacchi, 2012). Предполагается, что данные описываются уравнениями (3), но совместное распределение случайных ошибок неизвестно. На первом шаге оцениваются модели бинарной регрессии при помощи полупараметрической или непараметрической модели бинарного выбора. На втором шаге основное уравнение оценивается методом наименьших квадратов, при этом неизвестная функция g^* аппроксимируется с помощью сплайнов или полинома g_k , степень которого выбирается при помощи кросс-валидации:

$$g_k(X'_{i1} \hat{\beta}_1) = \sum_{m=1}^k \tau_m \lambda(X'_{i1} \hat{\beta}_1)^m, \quad (8)$$

где λ — обратное отношение Миллса, а математическое ожидание зависимой переменной имеет вид:

$$E(y_{2i}) = E(y_{2i}^* | y_{1i} = 1) = E(y_{2i}^* | y_{1i}^* \geq 0) = X'_{2i} \beta_2 + g^*(X'_{i1} \beta_1). \quad (9)$$

В работе значения линейного индекса $X'_{i1} \hat{\beta}_1$ на первом шаге оцениваются с помощью непараметрического метода, предложенного в (Klein, Spady, 1993).

Ниже последовательно представлены результаты оценивания вероятности оплаты занятий спортом с помощью пробит-моделей с учетом смещения выборки и величины оплаты занятий спортом с использованием моделей Хекмана, а также результаты оценивания величины оплаты занятий спортом с помощью полупараметрического метода Ньюи и обычного метода наименьших квадратов.

Для модели бинарного выбора и модели Хекмана гипотеза о равенстве коэффициентов в моделях, построенных на гендерных выборках, не была отвергнута, поэтому в работе эти модели рассматриваются на общей выборке. Для модели с полупараметрическим способом коррекции отбора приводится оценивание как на общей выборке, так и на подвыборках мужчин и женщин.

4.1. Пробит-модель с учетом смещения выборки

На первом этапе оценивание вероятности оплаты занятий спортом производится с помощью пробит-моделей с учетом смещения выборки — принятие решений о занятиях непрофессиональным спортом в оздоровительных целях оценивается по всей выборке. Затем на подвыборке из тех, кто занимается спортом, оценивается принятие решений об оплате занятий спортом. Гипотеза о независимости уравнения отбора и основного уравнения не отвергается только для первой спецификации, тогда как во второй и третьей решения индивида о занятии спортом и об оплате занятий спортом связаны. Результаты, полученные для разных

спецификаций и на разных подвыборках, похожи. Отличие состоит в том, что коэффициент при переменной логарифма дохода в основном уравнении оказался незначимым, тогда как для спецификаций с оценкой дохода через доход домохозяйств была обнаружена связь между доходами и вероятностью оплаты занятий спортом. Это может быть вызвано двумя причинами. Во-первых, респонденты при ответе на вопрос о размере доходов могли сознательно или несознательно исказить размер своих доходов, что приводило к ошибкам измерения. Во-вторых, 63.5% респондентов отказались называть свой доход, что привело к смещению подвыборки в сторону низкодоходных групп (при прямых ответах на вопросы о личных доходах индивиды, особенно располагающие высокими доходами, могут занижать размер своего дохода). Поэтому на используемых данных более корректно анализировать взаимосвязь дохода и занятий спортом при помощи переменных, характеризующих доход домохозяйств.

Полученные результаты свидетельствуют о том, что *вероятность оплаты занятий спортом в оздоровительных целях положительно связана с размером дохода домохозяйства*, в то время как связи дохода с вероятностью занятий спортом не обнаружено. Полученные значения коэффициентов регрессоров в уравнении отбора во всех спецификациях модели во многом соответствуют результатам зарубежных исследователей: женщины более склонны выбирать платные занятия спортом, если они решили заниматься спортом, а возраст индивида не связан с вероятностью оплаты занятий спортом. Положительная связь обнаружена между оплатой занятий спортом и такими факторами, как семейное положение индивида и занятия спортом вместе с близким окружением. Во всех спецификациях были значимыми коэффициенты при переменных, характеризующих место проживания: жители сельской местности выбирают занятия спортом, не требующие оплаты. Курение индивида, по-видимому, по-разному связано с вероятностью занятий спортом и вероятностью оплаты занятий спортом: в основном уравнении связь положительная, в уравнении отбора — отрицательная.

В связи с тем, что для пробит-модели можно интерпретировать только знаки коэффициентов, рассмотрим табл. 1, содержащую значимые предельные эффекты, которые показывают, насколько увеличивается или уменьшается вероятность занятий платными видами спорта у взрослого населения при изменении рассматриваемых в обоих уравнениях факторов.

Значения предельных эффектов, рассчитанные для вероятности оплаты занятий спортом с учетом смещения выборки, свидетельствуют о том, что вклад фактора доходов домохозяйства значим и составляет 6–7 процентных пунктов (пп.). Однако при построении модели на подвыборке из индивидов, ответивших на прямой вопрос о размере дохода, переменная дохода оказалась незначимой. Вероятность платных занятий спортом у женщин выше на 10–17 пп. по сравнению с мужчинами. Семейный статус ассоциируется с увеличением вероятности оплаты спортивных занятий на 8–12 пп. Проживание в сельской местности связано со снижением вероятности оплаты занятий спортом на 9–17 пп. Если друзья, коллеги по работе или учебе занимаются спортом, вероятность оплаты занятий спортом индивидом увеличивается на 12–13 пп.

Построение пробит-моделей с учетом смещения выборки позволило выявить, что некоторые независимые переменные (пол, курение) действуют разнонаправлено в уравнении отбора и в основном уравнении. Например, женщины меньше склонны заниматься спортом по сравнению с мужчинами, однако если они выбирают занятия спортом, то вероятность платить за них у женщин выше; а курение ассоциируется с оплатой занятий спортом, но отрицательно связано с их вероятностью.

Таблица 1. Предельные эффекты пробит-модели с учетом смещения выборки для вероятности выбора платных занятий спортом

	Спецификации модели		
	(1)	(2)	(3)
<i>sportpay</i>	<i>Основное уравнение</i>		
<i>male</i>	-0.175*** (0.047)	-0.110*** (0.030)	-0.096*** (0.028)
<i>civilstatus</i>	н/р	0.115*** (0.033)	0.079* (0.031)
<i>smoke</i>	н/р	0.088* (0.045)	0.077* (0.043)
<i>rural</i>	-0.171* (0.084)	-0.094* (0.042)	-0.099* (0.039)
<i>lnincome</i>	н/р	—	—
<i>lnahhinc</i>	—	0.0692* (0.0302)	—
<i>incomegroup</i>	—	—	0.060** (0.021)
<i>entourage</i>	н/р	0.119*** (0.032)	0.127*** (0.029)
<i>sport</i>	<i>Уравнение отбора</i>		
<i>male</i>	0.114*** (0.027)	0.084*** (0.018)	0.078*** (0.017)
<i>smoke</i>	-0.138*** (0.028)	-0.146*** (0.021)	-0.142*** (0.020)
Число наблюдений	1516	3278	3864

Примечание. В скобках приведены оценки стандартных отклонений; ***, **, * — значимость на уровне 1, 5 и 10% соответственно; «—» означает, что данная переменная не включалась в соответствующую модель; н/р — предельный эффект не рассчитывался, поскольку коэффициент при данной переменной был незначимым в рассматриваемой спецификации.

4.2. Модель Хекмана

В качестве уравнения отбора принимается решение индивида о занятиях спортом в оздоровительных целях, а в качестве основного уравнения — размер оплаты занятий спортом. Для оценивания используется модель Хекмана: модель оценивается методом максимального правдоподобия (MLE) и с помощью двухшаговой (two-step) процедуры (Heckman, 1976). Так же как и в п. 4.1, оцениваются три спецификации модели, по-разному учитывающие доходы индивида. Гипотеза о независимости уравнения отбора и основного уравнения отвергается, решения индивида об оплате занятий спортом и о размере оплаты занятий спортом связаны между собой. Коэффициент корреляции случайных ошибок имеет отрицательный знак, что говорит о том, что ненаблюдаемые факторы, увеличивающие склонность к оплате занятий спортом, уменьшают расходы на них.

Предельные эффекты полученных моделей представлены ниже в табл. 2, из которой видно, что регрессор дохода индивида оказался незначимым, но регрессоры дохода

домохозяйства и типа доходной группы домохозяйства значимы, а связь с размером оплаты занятий спортом — положительная. Предельный эффект гендерной принадлежности имеет отрицательный знак, т. е. женщины склонны больше платить за занятия спортом, чем мужчины. Гипотеза о связи размера оплаты занятий спортом с доходом индивида подтвердилась. Кроме того, положительная связь получена для размера оплаты занятий спортом с регрессорами вовлеченности окружения индивидов в занятия спортом и семейного статуса. Проживание в сельской местности отрицательно связано с размером оплаты занятий спортом.

Таблица 2. Предельные эффекты размера оплаты занятий спортом для модели Хекмана

	Спецификации модели					
	(1) MLE	(1) two-step	(2) MLE	(2) two-step	(3) MLE	(3) two-step
<i>male</i>	-1.181** (0.394)	-1.098** (0.395)	-0.609** (0.222)	-0.614** (0.233)	-0.557** (0.201)	-0.576** (0.209)
<i>civilstatus</i>	н/р	н/р	0.752** (0.230)	0.757** (0.234)	0.4438* (0.211)	0.458* (0.214)
<i>rural</i>	-1.177** (0.387)	-1.179** (0.385)	-0.634** (0.238)	-0.622* (0.242)	-0.653** (0.224)	-0.631** (0.230)
<i>lnincome</i>	н/р	н/р	—	—	—	—
<i>lnahhinc</i>	—	—	0.754*** (0.187)	0.747*** (0.189)	—	—
<i>incomegroup</i>	—	—	—	—	0.591*** (0.127)	0.585*** (0.130)
<i>entourage</i>	1.106*** (0.302)	1.131*** (0.304)	0.966*** (0.194)	0.948*** (0.197)	0.951*** (0.180)	0.932*** (0.182)
Число наблюдений	1516	1516	3278	3278	3864	3864

Примечание. В скобках приведены оценки стандартных отклонений; ***, **, * — значимость на уровне 1, 5 и 10% соответственно; «—» означает, что данная переменная не включалась в соответствующую модель; н/р — предельный эффект не рассчитывался, поскольку коэффициент при данной переменной был незначимым в рассматриваемой спецификации.

4.3. Модель с полупараметрическим способом коррекции отбора методом Ньюи на общей выборке

При моделировании принятия решений индивидов о занятиях спортом и размере их оплаты с помощью модели Хекмана предполагается, что распределение ошибок в уравнении участия и основном уравнении является нормальным. Ослабить это допущение позволяет применение полупараметрических методов моделирования. Для проверки полученных ранее результатов исследования были построены модели с полупараметрическим способом коррекции смещения отбора Ньюи (Newey, 2009), в которых в качестве основного уравнения выступала сумма затрат на занятия спортом, а в качестве уравнения участия — факт занятий спортом в оздоровительных целях. Для сравнения также были построены модели с помощью метода наименьших квадратов (OLS). Результаты представлены в табл. 3.

Таблица 3. Оценки размера оплаты занятий спортом с помощью полупараметрического метода Ньюи и метода наименьших квадратов на общей выборке

	Модели					
	(1) Newey	(1) OLS	(2) Newey	(2) OLS	(3) Newey	(3) OLS
<i>age2534</i>	0.228 (0.565)	0.464 (0.519)	0.989*** (0.289)	0.969*** (0.289)	0.696** (0.265)	0.716** (0.266)
<i>age3544</i>	-0.710 (0.705)	-0.282 (0.582)	0.563 (0.335)	0.402 (0.331)	0.620* (0.306)	0.485 (0.304)
<i>age4554</i>	-0.838 (0.648)	-0.492 (0.562)	0.092 (0.329)	-0.163 (0.316)	0.051 (0.306)	-0.179 (0.296)
<i>age5564</i>	-1.558* (0.691)	-1.234* (0.622)	-0.184 (0.301)	-0.429 (0.288)	-0.313 (0.286)	-0.529 (0.276)
<i>male</i>	-0.897* (0.36)	-0.943** (0.357)	-0.352 (0.21)	-0.235 (0.206)	-0.439* (0.199)	-0.263 (0.189)
<i>civilstatus</i>	0.528 (0.350)	0.569 (0.348)	0.521* (0.223)	0.367 (0.216)	0.270 (0.205)	0.119 (0.199)
<i>job1</i>	0.272 (0.326)	0.221 (0.323)	0.349 (0.243)	0.572* (0.229)	0.329 (0.217)	0.503* (0.208)
<i>hhsizе</i>	-0.088 (0.183)	-0.119 (0.181)	0.098 (0.117)	0.162 (0.115)	-0.050 (0.103)	0.024 (0.100)
<i>child</i>	-0.526 (0.401)	-0.504 (0.4)	-0.066 (0.296)	-0.099 (0.297)	-0.181 (0.266)	-0.286 (0.265)
<i>msc</i>	-0.625 (0.506)	-0.628 (0.506)	-0.168 (0.364)	-0.191 (0.365)	-0.007 (0.302)	0.015 (0.303)
<i>rural</i>	-1.2022** (0.392)	-1.2247** (0.391)	-0.7373** (0.234)	-0.7728*** (0.234)	-0.7577*** (0.221)	-0.8103*** (0.221)
<i>lnincome</i>	0.499 (0.265)	0.502 (0.265)	—	—	—	—
<i>lnahhinc</i>	—	—	0.8555*** (0.183)	0.8985*** (0.1829)	—	—
<i>incomegroup</i>	—	—	—	—	0.6925*** (0.1232)	0.7476*** (0.122)
<i>entourage</i>	1.238*** (0.304)	1.214*** (0.303)	1.114*** (0.195)	1.202*** (0.193)	1.054*** (0.181)	1.130*** (0.1796)
<i>smoke</i>	-0.407 (0.405)	-0.387 (0.404)	-0.154 (0.281)	-0.352 (0.272)	-0.037 (0.267)	-0.301 (0.251)
<i>drink</i>	0.057 (0.479)	0.081 (0.479)	-0.201 (0.331)	-0.234 (0.332)	0.139 (0.308)	0.074 (0.308)
<i>Intercept</i>	-3.427 (2.767)	-2.306 (2.562)	-5.0702* (2.068)	-7.868*** (1.791)	1.542 (0.956)	-0.934* (0.3897)
Adjusted R ²	0.0982	0.0979	0.1545	0.1492	0.1348	0.1297
Число наблюдений	483	483	1006	1006	1227	1227

Примечание. ***, **, * — значимость на уровне 1, 5 и 10% соответственно; « — » означает отсутствие регрессора в спецификации модели; для оценки с помощью метода Ньюи используется полином первой степени.

Из таблицы 3 видно, что направления влияния регрессоров для моделей OLS и Ньюи совпадают и соотносятся с предыдущими исследованиями. Статистически подтверждена гипотеза о положительной связи размера дохода индивида и размера затрат на спорт, что может быть объяснено с помощью эластичности по доходу — затраты на спорт по доходу являются товарами роскоши (Pawlowski, Breuer, 2012). Переменная, отвечающая за занятия спортом близкого окружения индивида, оказалась значимой во всех спецификациях, связь с переменной величины оплаты занятий спортом положительная. При рассмотрении моделей со спецификациями (2) и (3) установлено, что семейное положение, как и профессиональный статус, значимо связаны с расходами на спорт: руководители больше тратят на занятия спортом (Pawlowski, Breuer, 2012). Во всех моделях прослеживается отрицательная связь между проживанием в сельской местности и величиной оплаты занятий спортом: проживающие в сельской местности индивиды платят меньше.

Коэффициенты при фиктивных переменных, отвечающих за возраст индивида, свидетельствуют об обратной взаимосвязи затрат на спорт и возрастом индивида: молодые люди платят за спорт больше. Имеются существенные различия в затратах на спорт между мужчинами и женщинами: женщины склонны платить за занятия спортом больше, чем мужчины. В связи с этим построены модели на подвыборках по полу индивида.

4.4. Модель с полупараметрическим способом коррекции отбора методом Ньюи на подвыборках из мужчин и женщин

Направление влияния регрессоров для моделей OLS и Ньюи совпадают и согласуются с предыдущими исследованиями и при построении моделей на подвыборке из мужчин, результаты которых приведены в табл. 4. Согласно полученным результатам, для мужчин наблюдается отрицательная связь между затратами на спорт и курением, причиной которой, вероятно, является отрицательная связь между вероятностью занятий спортом и курением (Farrell, Shields, 2002; Downward et al., 2014). Наблюдается положительная взаимосвязь размера дохода домохозяйства и размера затрат на спорт для мужчин, но связи между личными доходами индивида и затратами на спорт не обнаружено. Совместные занятия спортом положительно связаны с размером затрат мужчин на спорт.

В таблице 5 представлены результаты оценивания затрат на спорт на подвыборке из женщин. Незначимость регрессоров перед членами полинома, корректирующего смещение отбора, говорит о том, что на подвыборке из женщин можно использовать обычную модель наименьших квадратов для оценивания размера оплаты занятий спортом. В то же время во всех спецификациях переменная, отвечающая за доход индивида, оказалась значимо связанной с размером оплаты занятий спортом.

Направление влияния остальных регрессоров для моделей OLS и Ньюи совпадают и соответствуют теоретическим предположениям. Семейный статус положительно связан с размером оплаты занятий спортом — замужние женщины тратят за занятия спортом больше, чем незамужние, также нахождение на должности руководителя или специалиста положительно связано с размером оплаты занятий спортом. В то время как в моделях на подвыборках из мужчин переменные возраста были не значимы, на подвыборках из женщин видна обратная связь возраста и затрат на спорт.

Таблица 4. Результаты регрессионного анализа оплаты занятий спортом с помощью полупараметрического метода Ньюи и метода наименьших квадратов на подвыборке из мужчин

	Модели					
	(1) Newey	(1) OLS	(2) Newey	(2) OLS	(3) Newey	(3) OLS
<i>age2534</i>	0.786 (0.717)	0.708 (0.667)	1.169* (0.482)	0.736 (0.458)	0.722 (0.440)	0.379 (0.411)
<i>age3544</i>	-0.302 (0.903)	-0.163 (0.846)	0.776 (0.632)	0.278 (0.55)	0.633 (0.542)	0.289 (0.493)
<i>age4554</i>	-0.505 (0.921)	-0.471 (0.849)	-0.060 (0.589)	-0.613 (0.538)	-0.083 (0.530)	-0.517 (0.500)
<i>age5564</i>	-1.074 (0.927)	-0.984 (0.900)	-0.379 (0.595)	-0.854 (0.532)	-0.549 (0.527)	-0.929 (0.504)
<i>civilstatus</i>	0.335 (0.561)	0.211 (0.564)	0.327 (0.384)	0.029 (0.366)	-0.021 (0.349)	-0.217 (0.332)
<i>job1</i>	0.223 (0.493)	0.199 (0.499)	0.395 (0.387)	0.463 (0.389)	0.451 (0.355)	0.476 (0.354)
<i>hhsizе</i>	-0.206 (0.243)	-0.193 (0.245)	-0.019 (0.192)	0.061 (0.191)	-0.071 (0.163)	0.005 (0.16)
<i>child</i>	-0.547 (0.642)	-0.649 (0.649)	-0.034 (0.503)	-0.019 (0.505)	-0.199 (0.447)	-0.189 (0.449)
<i>msc</i>	-0.149 (0.769)	0.154 (0.769)	0.347 (0.621)	0.246 (0.618)	0.439 (0.521)	0.319 (0.521)
<i>rural</i>	-0.649 (0.587)	-0.725 (0.593)	-0.485 (0.404)	-0.490 (0.407)	-0.482 (0.366)	-0.499 (0.368)
<i>lnincome</i>	-0.068 (0.395)	0.001 (0.396)	—	—	—	—
<i>lnahhinc</i>	—	—	0.732* (0.308)	0.704* (0.309)	—	—
<i>incomegroup</i>	—	—	—	—	0.587** (0.213)	0.672** (0.209)
<i>entourage</i>	1.291** (0.459)	1.2097** (0.463)	1.039** (0.329)	1.105*** (0.332)	0.925** (0.302)	0.937** (0.301)
<i>smoke</i>	-1.151* (0.486)	-1.016* (0.487)	-0.819* (0.352)	-0.849* (0.354)	-0.721* (0.329)	-0.797* (0.329)
<i>drink</i>	-0.099 (0.603)	-0.203 (0.605)	-0.304 (0.446)	-0.362 (0.448)	-0.0698 (0.399)	-0.119 (0.400)
<i>Intercept</i>	-64.765* (26.915)	2.258 (4.048)	3.263 (4.792)	-5.444 (3.105)	26.172* (12.147)	-0.356 (0.735)
Adjusted R^2	0.0805	0.0575	0.1261	0.1128	0.1114	0.1005
Число наблюдений	195	195	361	361	456	456

Примечание. ***, **, * — значимость на уровне 1, 5 и 10% соответственно; « — » означает отсутствие регрессора в спецификации модели; для оценки с помощью метода Ньюи используется полином третьей степени в спецификациях (1) и (3), в спецификации (2) используется полином второй степени.

Таблица 5. Результаты регрессионного анализа размера оплаты занятий спортом с помощью полупараметрического метода Ньюи и метода наименьших квадратов на подвыборке из женщин

	Модели					
	(1) Newey	(1) OLS	(2) Newey	(2) OLS	(3) Newey	(3) OLS
<i>age2534</i>	-0.167 (0.840)	-0.046 (0.828)	1.057** (0.385)	1.062** (0.384)	0.916* (0.358)	0.929** (0.357)
<i>age3544</i>	-1.156 (0.951)	-0.843 (0.875)	0.408 (0.429)	0.405 (0.429)	0.590 (0.396)	0.596 (0.396)
<i>age4554</i>	-1.014 (0.897)	-0.709 (0.821)	0.059 (0.401)	0.042 (0.395)	0.021 (0.371)	0.005 (0.37)
<i>age5564</i>	-1.888* (0.959)	-1.638 (0.912)	-0.215 (0.360)	-0.241 (0.346)	-0.286 (0.337)	-0.318 (0.332)
<i>civilstatus</i>	0.804 (0.479)	0.876 (0.472)	0.543 (0.284)	0.529 (0.279)	0.357 (0.266)	0.323 (0.258)
<i>job1</i>	0.112 (0.434)	0.116 (0.433)	0.599* (0.291)	0.597* (0.291)	0.439 (0.271)	0.474 (0.263)
<i>hhsizе</i>	-0.178 (0.278)	-0.194 (0.277)	0.144 (0.152)	0.152 (0.148)	-0.046 (0.132)	-0.038 (0.131)
<i>child</i>	0.047 (0.553)	0.021 (0.551)	0.019 (0.382)	0.015 (0.381)	-0.172 (0.340)	-0.179 (0.339)
<i>msc</i>	-1.096 (0.709)	-1.052 (0.707)	-0.293 (0.468)	-0.308 (0.464)	-0.151 (0.378)	-0.156 (0.378)
<i>rural</i>	-1.517** (0.526)	-1.514** (0.526)	-0.935** (0.289)	-0.938** (0.288)	-0.995*** (0.278)	-1.002*** (0.277)
<i>lnincome</i>	0.949* (0.368)	0.927* (0.366)	—	—	—	—
<i>lnahhinc</i>	—	—	0.932*** (0.231)	0.937*** (0.230)	—	—
<i>incomegroup</i>	—	—	—	—	0.731*** (0.152)	0.736*** (0.151)
<i>entourage</i>	1.285** (0.413)	1.249** (0.410)	1.244*** (0.244)	1.253*** (0.241)	1.184*** (0.229)	1.200*** (0.227)
<i>smoke</i>	0.793 (0.724)	0.789 (0.723)	0.399 (0.439)	0.392 (0.449)	0.489 (0.412)	0.461 (0.409)
<i>drink</i>	0.244 (0.795)	0.219 (0.794)	-0.064 (0.519)	-0.057 (0.518)	0.39 (0.501)	0.399 (0.501)
<i>Intercept</i>	-7.662* (3.847)	-6.272 (3.474)	-8.156** (2.506)	-8.444*** (2.246)	-0.515 (1.085)	-1.050* (0.474)
Adjusted R ²	0.1004	0.1013	0.1674	0.1686	0.1483	0.149
Число наблюдений	288	288	645	645	771	771

Примечание. ***, **, * — значимость на уровне 1, 5 и 10% соответственно; « — » означает отсутствие регрессора в спецификации модели; для оценки с помощью метода Ньюи используется полином первой степени.

5. Заключение

Целью работы было устранить пробел в исследовании принятия взрослыми индивидами решений об оплате занятий спортом и физкультурой, а также взаимосвязи этих решений с социально-демографическими и экономическими характеристиками индивидов. Для этого на российских данных рассматривались вероятностные модели бинарного выбора, модели Хекмана, а также модели с полупараметрическим способом коррекции смещения отбора, в целом, дающие похожие результаты. Оценивались три спецификации модели, отличающиеся друг от друга переменной, отвечающей за доход индивида, опираясь на результаты которых, можно сделать вывод о том, что для оценивания размера оплаты занятий спортом на рассматриваемых данных лучше всего использовать полупараметрический метод коррекции смещения отбора Ньюи, а среди рассмотренных переменных, отражающих доход индивида, использовать усредненный размер дохода домохозяйств. При построении моделей индивиды, занимающиеся спортом только профессионально, не рассматривались.

Расчитанные предельные эффекты моделей показали, что доход индивида положительно связан с вероятностью оплаты занятий спортом и размером их оплаты, как и в большинстве зарубежных работ. Использование разных переменных для измерения дохода было связано с их ограничениями (пропусками при ответах на вопросы о доходах) и с необходимостью проверить устойчивость результатов. Следует отметить, что результаты оценивания моделей разных спецификаций различаются, но незначительно.

Кроме того, были построены модели на подвыборках из мужчин и женщин для выявления особенностей их поведения. Для модели бинарного выбора и модели Хекмана гипотеза о равенстве коэффициентов в построенных на гендерных выборках моделях не была отвергнута, поэтому в работе эти модели рассматривались на общей выборке. Для модели с полупараметрическим способом коррекции отбора оценивание проводилось как на общей выборке, так и на подвыборках мужчин и женщин. Оценивались вероятность оплаты занятий спортом и размер оплаты. Результаты показали, что неслучайный отбор имел место не среди тех индивидов, кто платит или не платит за занятия спортом, а среди тех, кто занимается или не занимается спортом. Полученные оценки говорят об устойчивости построенных моделей, а также о наличии существенных различий в оплате занятий спортом для мужчин и женщин.

Женщины менее склонны заниматься спортом по сравнению с мужчинами, однако если занимаются, то вероятность платить за занятия у них выше, как и размер оплаты. Также для женщин нахождение на должности руководителя или специалиста положительно связано с размером оплаты занятий спортом. Характерной особенностью для мужчин является наличие отрицательной связи между затратами на спорт и курением. В целом результаты по характеризующим вредные привычки факторам совпадают с зарубежными исследованиями: курение как минимум 1 раз в день отрицательно связано с вероятностью занятий спортом.

Среди социально-демографических факторов три следующих оказались значимыми в большинстве спецификаций: статус состояния в браке (положительно связан с размером оплаты занятий спортом), проживание в сельской местности (склонны платить за занятия спортом меньше), занятия спортом вместе с членами семьи, друзьями, коллегами (связаны с увеличением расходов на спорт).

Таким образом, в результате проведенного эмпирического анализа были выделены склонные к затратам на спорт группы населения и готовые платить больше за занятия спортом.

Во-первых, совместные занятия спортом с членами семьи, друзьями, коллегами положительно связаны с оплатой и размером оплаты занятий спортом — при разработке стратегии развития физкультуры и спорта стоит обращать внимание на возможность совместного времяпрепровождения и укрепления здоровья. Во-вторых, женщины менее вероятно занимаются спортом и, если занимаются, склонны больше платить за занятия спортом. Наконец, еще одна группа, которая заслуживает внимания — это курящие индивиды, которые реже, чем другие индивиды, занимаются спортом в оздоровительных целях и, касаясь мужчин, тратят меньше на занятия спортом.

Благодарности. Автор выражает признательность Б. Потанину за ценные замечания и рекомендации при построении полупараметрических моделей, а также анонимному рецензенту за высказанные предложения по выбору эконометрической модели.

Список литературы

Засимова Л. С., Локтев Д. А. (2016). Занятия спортом — удел богатых? (Эмпирический анализ занятий спортом в России). *Экономический журнал Высшей школы экономики*, 20 (3), 471–499.

Шишкин С. В., Власов В. В., Боярский С. Г., Засимова Л. С., Колосницына М. Г., Кузнецов П. П., Овчарова Л. Н., Сажина С. В., Степанов И. М., Хоркина Н. А., Шевский В. И., Шейман И. М., Якобсон Л. И. (2017). *Здравоохранение: современное состояние и возможные сценарии развития: доклад*. М.: Изд-во НИУ ВШЭ.

Alexandris K., Carroll B. (1999). Constraints on recreational sport participation in adults in Greece: Implications for providing and managing sport services. *Journal of Sport Management*, 13 (4), 317–332.

Andreff W. (2001). The correlation between economic underdevelopment and sport. *European Sport Management Quarterly*, 1 (4), 251–279.

Barber N., Havitz M. E. (2001). Canadian participation rates in ten sport and fitness activities. *Journal of Sport Management*, 15 (1), 51–76.

Breuer C., Wicker P. (2008). Demographic and economic factors influencing inclusion in the German sport system — a microanalysis of the years 1985 to 2005. *European Journal for Sport and Society*, 5 (1), 33–42.

Breuer C., Wicker P. (2009). Decreasing sports activity with increasing age? Findings from a 20-year longitudinal and cohort sequence analysis. *Research Quarterly for Exercise and Sport*, 80 (1), 22–31.

Breuer C., Hallmann K., Wicker P., Feiler S. (2010). Socio-economic patterns of sport demand and ageing. *European Review of Aging and Physical Activity*, 7 (2), 61–70.

Caruso R. (2011). Crime and sport participation: Evidence from Italian regions over the period 1997–2003. *The Journal of Socio-Economics*, 40 (5), 455–463.

Cawley J. (2004). An economic framework for understanding physical activity and eating behaviors. *American Journal of Preventive Medicine*, 27 (3), 117–125.

Cicchetti C. J., Seneca J. J., Davidson P. (1969). *The demand and supply of outdoor recreation: An economic analysis*. New Brunswick, N. J.: Bureau of Economic Research, Rutgers — The State University.

De Luca G., Peracchi F. (2012). Estimating Engel curves under unit and item nonresponse. *Journal of Applied Econometrics*, 27 (7), 1076–1099.

Downward P. (2004). On leisure demand: A post Keynesian critique of neoclassical theory. *Journal of Post Keynesian Economics*, 26 (3), 371–394.

- Downward P. (2007). Exploring the economic choice to participate in sport: Results from the 2002 General Household Survey. *International Review of Applied Economics*, 21 (5), 633–653.
- Downward P., Rasciute S. (2010). The relative demands for sports and leisure in England. *European Sport Management Quarterly*, 10 (2), 189–214.
- Downward P., Rasciute S. (2011). Does sport make you happy? An analysis of the well-being derived from sports participation. *International Review of Applied Economics*, 25 (3), 331–348.
- Downward P., Riordan J. (2007). Social interactions and the demand for sport: An economic analysis. *Contemporary Economic Policy*, 25 (4), 518–537.
- Downward P., Lera-Lopez F., Rasciute S. (2011). The zero-inflated ordered probit approach to modeling sports participation. *Economic Modelling*, 28 (6), 2469–2477.
- Downward P., Lera-Lopez F., Rasciute S. (2014). The correlates of sports participation in Europe. *European Journal of Sport Science*, 14 (6), 592–602.
- Farrell L., Shields M. A. (2002). Investigating the economic and demographic determinants of sporting participation in England. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 165 (2), 335–348.
- García J., Lera-López F., Suárez M. J. (2011). Estimation of a structural model of the determinants of the time spent on physical activity and sport: Evidence for Spain. *Journal of Sports Economics*, 12 (5), 515–537.
- Heckman J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic Social Measurement*, 5 (4), 475–492.
- Humphreys B. R., Ruseski J. E. (2006). Economic determinants of participation in physical activity and sport. *Working Paper Series* 06–13. International Association of Sports Economists.
- Humphreys B. R., Ruseski J. E. (2007). Participation in physical activity and government spending on parks and recreation. *Contemporary Economic Policy*, 25 (4), 538–552.
- Klein R., Spady R. (1993). An efficient semiparametric estimator for binary response models. *Econometrica*, 61 (2), 387–421.
- Lechner M. (2009). Long-run labour market and health effects of individual sports activities. *Journal of Health Economics*, 28 (4), 839–854.
- Lera-López F., Rapún-Gárate M. (2007). The demand for sport: Sport consumption and participation models. *Journal of Sport Management*, 21 (1), 103–122.
- Miles L. (2007). Physical activity and health. *Nutrition Bulletin*, 32 (4), 314–363.
- Moens M., Scheerder J. (2004). Social determinants of sports participation revisited. The role of socialization and symbolic trajectories. *European Journal for Sport and Society*, 1 (1), 35–49.
- Newey W. K. (2009). Two-step series estimation of sample selection models. *The Econometrics Journal*, 12, 217–229.
- Pawlowski T., Breuer C. (2011). The demand for sports and recreational services: Empirical evidence from Germany. *European Sport Management Quarterly*, 11 (1), 5–34.
- Pawlowski T., Breuer C. (2012). Expenditure elasticities of the demand for leisure services. *Applied Economics*, 44 (26), 3461–3477.
- Pfeifer C., Cornelißen T. (2010). The impact of participation in sports on educational attainment — New evidence from Germany. *Economics of Education Review*, 29 (1), 94–103.
- Scheerder J., Vanreusel B., Taks M. (2005). Stratification patterns of active sport involvement among adults: Social change and persistence. *International Review for the Sociology of Sport*, 40 (2), 139–162.

Scheerder J., Vos S., Taks M. (2011). Expenditures on sport apparel: Creating consumer profiles through interval regression modelling. *European Sport Management Quarterly*, 11 (3), 251–274.

Stempel C. (2005). Adult participation sports as cultural capital: A test of Bourdieu's theory of the field of sports. *International Review for the Sociology of Sport*, 40 (4), 411–432.

Strawinski P. (2010). Economic determinants of sport participation in Poland. *Rivista di Diritto ed Economia dello Sport*, 3, 55–76.

Thibaut E., Vos S., Scheerder J. (2014). Hurdles for sports consumption? The determining factors of household sports expenditures. *Sport Management Review*, 17 (4), 444–454.

WHO. (2010). *Global recommendations on physical activity for health*. World Health Organization. <https://www.who.int/dietphysicalactivity/publications/9789241599979/en/>.

Wicker P., Breuer C., Pawlowski T. (2009). Promoting sport for all to age-specific target groups: The impact of sport infrastructure. *European Sport Management Quarterly*, 9 (2), 103–118.

Wicker P., Breuer C., Pawlowski T. (2010). Are sports club members big spenders?: Findings from sport specific analyses in Germany. *Sport Management Review*, 13 (3), 214–224.

Wilson T. C. (2002). The paradox of social class and sports involvement: The roles of cultural and economic capital. *International Review for the Sociology of Sport*, 37 (1), 5–16.

Поступила в редакцию 25.05.2019;
принята в печать 27.09.2020.

Приложение

Таблица III. Основные описательные статистики непрерывных переменных

Переменная	Описание	Число наблюдений	Среднее	Стандартное отклонение	Минимум	Максимум
<i>age</i>	Возраст	4006	47.56	17.71	16	97
<i>agesq</i>	Квадрат возраста	4006	2575.52	1768.62	256	9409
<i>income</i>	Доходы индивида	1569	23836.86	15145.19	500	120000
<i>lnincome</i>	Логарифм доходов	1569	9.88	0.67	6.21	11.69
<i>ahhinc</i>	Доходы на одного члена домохозяйства	3375	16245.34	10467.54	625	95000
<i>lnahhinc</i>	Логарифм доходов на одного члена домохозяйства	3375	9.52	0.6	6.43	11.46
<i>sportpaysum</i>	Расходы на спорт	1229	517.97	1279.4	0	15000
<i>lnsps</i>	Логарифм расходов на спорт	1229	1.83	3.19	0	9.62

Таблица П2. Основные описательные статистики дискретных переменных

Переменная	Описание	Число наблюдений
<i>sport</i> — занятие спортом в оздоровительных целях	1 — занимается спортом;	1229
	0 — иначе	2644
<i>sportpay</i> — оплата занятий спортом	1 — занимается спортом и платит за занятия спортом;	309
	0 — занимается спортом и не платит за занятия спортом	920
<i>male</i> — пол	1 — мужской;	1390
	0 — женский	2616
<i>infrastructure</i> — наличие спортивной инфраструктуры по месту работы/учебы	1 — спортивная инфраструктура есть;	2369
	0 — иначе	1637
<i>entourage</i> — наличие занимающихся спортом вместе с индивидом в окружении индивида	1 — члены семьи, друзья, коллеги занимаются спортом вместе с индивидом;	496
	0 — иначе	3510
<i>bhealth</i> — плохое здоровье	1 — плохое здоровье;	739
	0 — иначе	3267
<i>nhealth</i> — среднее здоровье	1 — среднее здоровье;	2371
	0 — иначе	1635
<i>ghealth</i> — хорошее здоровье	1 — хорошее здоровье;	1518
	0 — иначе	2488
<i>high_educ</i> — высшее образование	1 — высшее образование;	1265
	0 — иначе	2741
<i>civilstatus</i> — официальный или гражданский брак	1 — состоит в браке;	2070
	0 — иначе	1936
<i>married</i> — официальный брак	1 — состоит в официальном браке;	1908
	0 — иначе	2098
<i>child</i> — наличие детей	1 — есть дети моложе 18 лет;	1220
	0 — иначе	2786
<i>incomegroup</i> — доходная группа индивида	2 — с низким или ниже среднего доходом;	1207
	3 — со средним уровнем дохода;	1934
	4 — доход выше среднего;	772
	5 — с высоким или самым высоким уровнем дохода	83
<i>smoke</i> — курение	1 — курит (как минимум 1 раз в день);	920
	0 — иначе	3086
<i>drink</i> — алкоголь	1 — употребляет алкогольные напитки чаще, чем 2 раза в неделю;	430
	0 — иначе	3576
<i>job1</i> — руководители, специалисты	1 — руководитель, специалист;	891
	0 — иначе	3115
<i>job2</i> — квалифицированные рабочие	1 — квалифицированный рабочий;	538
	0 — иначе	3468
<i>job3</i> — неквалифицированные рабочие	1 — неквалифицированный рабочий;	131
	0 — иначе	3875
<i>job4</i> — «силовики»	1 — военнослужащий в армии, работник милиции, ГИБДД, органов безопасности;	21
	0 — иначе	3985

Окончание табл. П2

Переменная	Описание	Число наблюдений
<i>rural</i> — сельская местность	1 — проживает в сельской местности; 0 — иначе	1003 3003
<i>msc</i> — Москва	1 — проживает в Москве; 0 — иначе	342 3664
<i>fo1</i> — Северо-Западный ФО	1 — проживает в Северо-Западном ФО; 0 — иначе	389 3617
<i>fo2</i> — Центральный ФО	1 — проживает в Центральном ФО; 0 — иначе	1092 2914
<i>fo3</i> — Южный ФО	1 — проживает в Южном ФО; 0 — иначе	689 3317
<i>fo4</i> — Поволжский ФО	1 — проживает в Поволжском ФО; 0 — иначе	817 3189
<i>fo5</i> — Уральский ФО	1 — проживает в Уральском ФО; 0 — иначе	332 3674
<i>fo6</i> — Сибирский ФО	1 — проживает в Сибирском ФО; 0 — иначе	518 3488
<i>fo7</i> — Дальневосточный ФО	1 — проживает в Дальневосточном ФО; 0 — иначе	169 3837

Makshanchikov K. Russians' spending on sports: Econometric analysis on Levada-Center data. *Applied Econometrics*, 2020, v. 60, pp. 115–138.

DOI: 10.22394/1993-7601-2020-60-115-138

Konstantin Makshanchikov

National Research University Higher School of Economics, Moscow, Russian Federation;
kmakehanchikov@hse.ru

Russians' spending on sports: Econometric analysis on Levada-Center data

This study aims to examine how Russian adults make their choices about expenditures on non-professional sports to improve their health. The data was taken from Levada-Center survey on the attitude of people to their health and the quality of medical care in Russia, conducted in 2017. Probabilistic models of binary choice, Heckman selection model and semiparametric model by Newey were employed. The results of the study showed positive relationship between expenditures on sports and individual's income. Gender, age, profession, rural and entourage were among other factors that determined individual expenditures on sport.

Keywords: sport; expenditures on sport; human capital; healthy lifestyle; non-professional sport.

JEL classification: C21; C31; I12; Z20.

References

Zasimova L. S., Loktev D. A. (2016). Sports for the rich? (Empirical investigation of participation in sport in Russia). *The HSE Economic Journal*, 20, 3, 471–498 (in Russian).

Shishkin S. V., Vlasov V. V., Boyarskiy S. G., Zasimova L. S., Kolosnitsyna M. G., Kuznetsov P. P., Ovcharova L. N., Sazhina S. V., Stepanov I. M., Khorkina N. A., Shevskiy V. I., Sheyman I. M., Yakobson L. I. (2017). Healthcare: current status and possible development scenarios: report. Moscow: Higher School of Economics (in Russian).

Alexandris K., Carroll B. (1999). Constraints on recreational sport participation in adults in Greece: Implications for providing and managing sport services. *Journal of Sport Management*, 13 (4), 317–332.

Andreff W. (2001). The correlation between economic underdevelopment and sport. *European Sport Management Quarterly*, 1 (4), 251–279.

Barber N., Havitz M. E. (2001). Canadian participation rates in ten sport and fitness activities. *Journal of Sport Management*, 15 (1), 51–76.

Breuer C., Wicker P. (2008). Demographic and economic factors influencing inclusion in the German sport system — a microanalysis of the years 1985 to 2005. *European Journal for Sport and Society*, 5 (1), 33–42.

Breuer C., Wicker P. (2009). Decreasing sports activity with increasing age? Findings from a 20-year longitudinal and cohort sequence analysis. *Research Quarterly for Exercise and Sport*, 80 (1), 22–31.

Breuer C., Hallmann K., Wicker P., Feiler S. (2010). Socio-economic patterns of sport demand and ageing. *European Review of Aging and Physical Activity*, 7 (2), 61–70.

Caruso R. (2011). Crime and sport participation: Evidence from Italian regions over the period 1997–2003. *The Journal of Socio-Economics*, 40 (5), 455–463.

Cawley J. (2004). An economic framework for understanding physical activity and eating behaviors. *American Journal of Preventive Medicine*, 27 (3), 117–125.

Cicchetti C. J., Seneca J. J., Davidson P. (1969). *The demand and supply of outdoor recreation: An econometric analysis*. New Brunswick, N. J.: Bureau of Economic Research, Rutgers — The State University.

De Luca G., Peracchi F. (2012). Estimating Engel curves under unit and item nonresponse. *Journal of Applied Econometrics*, 27 (7), 1076–1099.

Downward P. (2004). On leisure demand: A post Keynesian critique of neoclassical theory. *Journal of Post Keynesian Economics*, 26 (3), 371–394.

Downward P. (2007). Exploring the economic choice to participate in sport: Results from the 2002 General Household Survey. *International Review of Applied Economics*, 21 (5), 633–653.

Downward P., Rasciute S. (2010). The relative demands for sports and leisure in England. *European Sport Management Quarterly*, 10 (2), 189–214.

Downward P., Rasciute S. (2011). Does sport make you happy? An analysis of the well-being derived from sports participation. *International Review of Applied Economics*, 25 (3), 331–348.

Downward P., Riordan J. (2007). Social interactions and the demand for sport: An economic analysis. *Contemporary Economic Policy*, 25 (4), 518–537.

Downward P., Lera-Lopez F., Rasciute S. (2011). The zero-inflated ordered probit approach to modelling sports participation. *Economic Modelling*, 28 (6), 2469–2477.

Downward P., Lera-Lopez F., Rasciute S. (2014). The correlates of sports participation in Europe. *European Journal of Sport Science*, 14 (6), 592–602.

Farrell L., Shields M. A. (2002). Investigating the economic and demographic determinants of sporting participation in England. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 165 (2), 335–348.

García J., Lera-López F., Suárez M. J. (2011). Estimation of a structural model of the determinants of the time spent on physical activity and sport: Evidence for Spain. *Journal of Sports Economics*, 12 (5), 515–537.

Heckman J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic Social Measurement*, 5 (4), 475–492.

Humphreys B. R., Ruseski J. E. (2006). Economic determinants of participation in physical activity and sport. *Working Paper Series* 06–13. International Association of Sports Economists.

Humphreys B. R., Ruseski J. E. (2007). Participation in physical activity and government spending on parks and recreation. *Contemporary Economic Policy*, 25 (4), 538–552.

Klein R., Spady R. (1993). An efficient semiparametric estimator for binary response models. *Econometrica*, 61 (2), 387–421.

Lechner M. (2009). Long-run labour market and health effects of individual sports activities. *Journal of Health Economics*, 28 (4), 839–854.

Lera-López F., Rapún-Gárate M. (2007). The demand for sport: Sport consumption and participation models. *Journal of Sport Management*, 21 (1), 103–122.

Miles L. (2007). Physical activity and health. *Nutrition Bulletin*, 32 (4), 314–363.

Moens M., Scheerder J. (2004). Social determinants of sports participation revisited. The role of socialization and symbolic trajectories. *European Journal for Sport and Society*, 1 (1), 35–49.

Newey W. K. (2009). Two-step series estimation of sample selection models. *The Econometrics Journal*, 12, 217–229.

Pawlowski T., Breuer C. (2011). The demand for sports and recreational services: Empirical evidence from Germany. *European Sport Management Quarterly*, 11 (1), 5–34.

Pawlowski T., Breuer C. (2012). Expenditure elasticities of the demand for leisure services. *Applied Economics*, 44 (26), 3461–3477.

Pfeifer C., Cornelißen T. (2010). The impact of participation in sports on educational attainment — New evidence from Germany. *Economics of Education Review*, 29 (1), 94–103.

Scheerder J., Vanreusel B., Taks M. (2005). Stratification patterns of active sport involvement among adults: Social change and persistence. *International Review for the Sociology of Sport*, 40 (2), 139–162.

Scheerder J., Vos S., Taks M. (2011). Expenditures on sport apparel: Creating consumer profiles through interval regression modelling. *European Sport Management Quarterly*, 11 (3), 251–274.

Stempel C. (2005). Adult participation sports as cultural capital: A test of Bourdieu's theory of the field of sports. *International Review for the Sociology of Sport*, 40 (4), 411–432.

Strawinski P. (2010). Economic determinants of sport participation in Poland. *Rivista di Diritto ed Economia dello Sport*, 3, 55–76.

Thibaut E., Vos S., Scheerder J. (2014). Hurdles for sports consumption? The determining factors of household sports expenditures. *Sport Management Review*, 17 (4), 444–454.

WHO. (2010). *Global recommendations on physical activity for health*. World Health Organization. <https://www.who.int/dietphysicalactivity/publications/9789241599979/en/>.

Wicker P., Breuer C., Pawlowski T. (2009). Promoting sport for all to age-specific target groups: The impact of sport infrastructure. *European Sport Management Quarterly*, 9 (2), 103–118.

Wicker P., Breuer C., Pawlowski T. (2010). Are sports club members big spenders?: Findings from sport specific analyses in Germany. *Sport Management Review*, 13 (3), 214–224.

Wilson T. C. (2002). The paradox of social class and sports involvement: The roles of cultural and economic capital. *International Review for the Sociology of Sport*, 37 (1), 5–16.

Received 25.05.2019; accepted 27.09.2020.