

ОТРАСЛЕВАЯ И РЕГИОНАЛЬНАЯ ЭКОНОМИКА

Ф. И. Миничев¹

ЦМАКП / ИНП РАН / НИУ ВШЭ / МГУ имени М. В. Ломоносова
(Москва, Россия)

О. А. Маркова²

МГУ имени М. В. Ломоносова / РАНХиГС (Москва, Россия)

УДК: 334.01, 334.021

doi: 10.55959/MSU0130-0105-6-59-2-9

ФАКТОРЫ ВХОДА НА РЫНКИ ПАССАЖИРСКИХ АВИАПЕРЕВОЗОК

Отрасль пассажирских авиаперевозок динамично развивается, и, хотя появление авиакомпаний — редкое явление, существующие авиакомпании активно конкурируют на маршрутах. Вход фирмы на новое направление может говорить о его прибыльности, которую нельзя наблюдать напрямую, поэтому модели входа могут дать дополнительную информацию о рынке, которую иначе получить сложно. В исследовании используются данные о перелетах внутри России, характеристиках маршрутов и характеристиках присутствия авиакомпаний на рынках, чтобы выявить факторы, которые влияют на вероятность авиакомпании войти на маршрут. На основе этого дается тестирование модели входа, предложенная в работе (Berry, 1992). Для выявления факторов, влияющих на вход авиакомпании на маршрут между двумя городами внутри России, используется модель пробит. Далее показана устойчивость полученных результатов: полученные оценки с использованием алгоритмов машинного обучения для решения задачи бинарного выбора не отличаются значительно от полученных в модели пробит оценок. Полученные результаты частично согласуются с результатами предшествующих исследований: население и количество туристических городов на маршруте (прокси для спроса на рынке) значимо положительно влияют на вероятность входа авиакомпаний. Также значимое положительное влияние на вход оказывает переменная, характеризующая присутствие авиакомпаний в аэропортах маршрута в предыдущем периоде. В отличие от исследований на европейских и американских данных, в данной работе показано отрицательное влияние расстояния на вероятность входа, что может быть объяснено особенностями российской географии. Этот вывод важен для проведения эффективной политики в области туризма и развития пассажирских перевозок.

¹ Миничев Федор Ильич — лаборант-исследователь, Центр макроэкономического анализа и краткосрочного прогнозирования; лаборант, Институт народнохозяйственного прогнозирования РАН; стажер-исследователь, Национальный исследовательский университет «Высшая школа экономики»; инженер, Экономический факультет МГУ имени М. В. Ломоносова; e-mail: fedor.minichev@gmail.com, ORCID: 0009-0007-9827-2758.

² Маркова Ольга Анатольевна — ассистент, Экономический факультет МГУ имени М. В. Ломоносова; научный сотрудник, РАНХиГС; e-mail: olya.eule@gmail.com, ORCID: 0000-0002-3512-6282.

© Миничев Федор Ильич, 2024 

© Маркова Ольга Анатольевна, 2024 

Ключевые слова: модели входа, авиакомпании, конкуренция, сетевые модели авиаперевозок.

Цитировать статью: Миничев, Ф. И., & Маркова, О. А. (2024). Факторы входа на рынки пассажирских авиаперевозок. *Вестник Московского университета. Серия 6. Экономика*, 59(2), 177–204. <https://doi.org/10.55959/MSU0130-0105-6-59-2-9>.

F. I. Minichev

CMASF / IEF RAS / HSE University / Lomonosov Moscow State University (Moscow, Russia)

O. A. Markova

Lomonosov Moscow State University / RANEPA (Moscow, Russia)

JEL: D22, L13, L93

DETERMINANTS OF ENTRY IN THE AIRLINE MARKETS

Air passenger transportation industry is dynamically expanding, and although the emergence of new airline companies is rare, existing companies are actively competing for routes. Firm's entry into a new destination can speak for its profitability, which cannot be directly observed, so entry patterns can provide additional information about the market that is otherwise difficult to obtain. The study is based on Russian domestic flights data, route characteristics, and airline market presence characteristics to identify factors that may influence airline's likelihood of entering a route. Using this dataset, we estimate the entry model proposed earlier (Berry, 1992). To identify the factors affecting airline entry on a route between two cities within Russia, we use probit. To validate our results we use machine learning algorithms to solve a binary choice problem. The results obtained are partly consistent with the results of previous studies: there is a significant positive effect of population and the number of tourist cities on the route (proxies for demand). Also, a variable characterizing the presence of the airline at the airports of the route in the previous period has a significant positive effect on the probability of airline enter a route. Unlike the studies on European and US data, our paper shows a negative effect of distance on entry probability, which can be explained by the peculiarities of Russian geography. This conclusion is important for effective policy making in Russian regions especially in promoting tourism and developing passenger transportation.

Keywords: airline companies, entry models, competition, airline network models.

To cite this document: Minichev, F. I., & Markova, O. A. (2024). Determinants of entry on airline markets. *Lomonosov Economics Journal*, 59(2), 177–204. <https://doi.org/10.55959/MSU0130-0105-6-59-2-9>

Введение

В данной работе мы изучаем факторы входа на рынки пассажирских авиаперевозок. Для анализа структуры рынка и условий конкуренции исследователю, как правило, требуются данные о ценах, продажах и издерж-

ках компаний, получить доступ к которым чаще всего затруднительно. Модели входа, ставшие важной областью исследований в современных эмпирических отраслевых рынках (Berry, Reiss 2007; Einav, Levin, 2010; Aguirregabiria, Но, 2012), также дают возможность делать выводы о характере конкуренции на рынке. При этом, во-первых, модели входа менее требовательные к данным по сравнению с другими подходами (оценкой спроса, исследованиями производственной функции и т.д.), а во-вторых, на основе предположения о том, что факт входа компании на рынки выявляет предпочтения, исследователь может судить о существующих издержках входа на рынок. Оценка таких моделей позволяет понять, как характеристики рынка влияют на вход нового игрока и как вход нового игрока воздействует на прибыльность остальных игроков.

В данном исследовании оценка модели входа проведена на российских данных о пассажирских авиаперевозках по следующим причинам. Во-первых, наибольшее развитие модели входа получили в исследованиях рынков пассажирских авиаперевозок, однако большая часть эмпирических исследований посвящена США, в том числе по причине легкости сбора данных. Исследование на российских данных представляет интерес, так как российские рынки пассажирских перевозок могут значительно отличаться от рынков США (в первую очередь из-за особенностей географии и расселения людей по территории страны). Во-вторых, в данном контексте рынок — ненаправленный¹ маршрут между двумя городами (маршруты АВ и ВА приравниваются друг к другу), что дает исследователю возможность рассматривать множество рынков внутри одной страны или региона. В-третьих, сектор авиаперевозок является важной частью экономики, характеризующей как технологический прогресс, так и мобильность и платежеспособность населения. Тем не менее конкуренция на российских рынках пассажирских авиаперевозок активно обсуждается (в части спроса на пассажирские авиаперевозки (Комаристый, 2004), существующих барьеров входа, которые выделяют участник отраслей (Лукьянов и др., 2007; Lukyanov et al., 2009, 2018)). В то время как некоторые исследователи предлагают математические модели (см. например, (Матюха, 2022)), модель (Berry, 1992) и ее расширения сравнительно более строгие и одновременно дают возможность состыковать предпосылки, теоретические выводы и эмпирические данные. Для рынков российских авиаперевозок модель (Berry, 1992) не оценивалась, поэтому в данном исследовании мы тестируем ее на российских данных.

¹ Хотя некоторые исследователи и называют рынком пассажирских перевозок все перелеты внутри страны, по факту для потребителей взаимозаменяемыми выступают маршруты между двумя городами. Этот подход используется во всех исследованиях в новых эмпирических отраслевых рынках (см.: (Aguirregabiria, 2012, p. 175)).

Рынок авиаперевозок, динамично развивавшийся до начала 2020 г.¹, в последние годы сталкивался с событиями, сильно замедляющими его развитие². Ограничительные меры в период пандемии COVID-19, напряженная обстановка в мире и санкции западных стран после 24 февраля 2022 г. привели к тому, что потенциальные маршруты авиалиний сузились до преимущественно российских, на которые раньше авиалинии могли меньше входить из-за наличия других маршрутов³. Кроме того, самолеты российских авиалиний либо иностранного производства, либо в их производстве используются иностранные компоненты, поэтому в условиях санкций авиакомпаниям особенно важно рационально распоряжаться имеющимися самолетами. В ответ на эти события правительство объявляет о поддержке гражданской авиации⁴. Для оказания эффективной поддержки необходимо понимать, как устроена конкуренция в отрасли и какие факторы влияют на вход авиалинии на новый маршрут.

Существует много исследований факторов входа на новые маршруты, но эти исследования проводились преимущественно на американских данных (Berry, 1992; Ciliberto, Tamer, 2009 и др.), и реже — на европейских (Bontemps, Sampaio, 2020). Таким образом, хотя существует множество оценок моделей входа на авиарынки, ни одно исследование еще не было проведено для России. А между тем, полученные в такого рода моделях результаты могут быть использованы для корректировки регулирования в отношении таких рынков.

Таким образом, исследовательский вопрос, поставленный в данной работе, — какие факторы влияют на вероятность входа фирмы на маршрут внутри России.

В своем исследовании мы собрали авторский датасет об авиаперелетах внутри России. На этих данных мы оценили пробит-модель вероятности входа авиакомпаний на маршруты, а затем, используя модели машинного обучения, подтвердили полученный результат для выявленных факторов, влияющих на вероятность входа. Исходя из полученных оценок, можно

¹ Годовой отчет компании Аэрофлот за 2019 год// Официальный сайт компании Аэрофлот. URL: <http://ar2019.aeroflot.ru/reports/aeroflot/annual/2019/gb/Russian/2020/.html> (дата обращения: 20.09.2023)

² Взлетел ли авиарынок?// Аналитический комментарий АКРА. URL: <https://www.asga-ratings.ru/research/2674/> (дата обращения: 20.09.2023).

³ Сокращение числа возможных маршрутов снизило альтернативные издержки входа на российские маршруты. Кроме того, в результате ограничений на международные перелеты могли произойти структурные сдвиги в спросе, что также могло повлиять на решения авиакомпаний. В настоящее время сложно судить о наличии равновесия на рассматриваемых рынках (поэтому в качестве рассматриваемого периода мы выбрали период до начала пандемии).

⁴ Правительство предоставит субсидии аэропортам и авиакомпаниям// РИА новости. URL: <https://ria.ru/20220407/subsidii-1782286064.html> (дата обращения: 20.09.2023).

говорить о значимом положительном влиянии на вероятность входа населения (прокси для спроса на рынке) и переменной, характеризующей присутствие авиакомпании в аэропортах маршрута в предыдущем периоде. Эти результаты, подтвержденные построением моделей машинного обучения, согласуются с предыдущими исследованиями ((Berry, 1992) и др.). В отличие от исследований на европейских и американских данных показано отрицательное влияние расстояния на вероятность входа, что может быть вызвано особенностями российской географии (большая протяженность с запада на восток).

Дальнейшая часть работы устроена следующим образом. В обзоре литературы мы рассматриваем особенности, свойственные рынкам авиаперевозок, приводим теоретические аспекты моделей входа и результаты их эмпирических оценок. Затем описаны методы сбора данных и приведен их визуальный анализ. В четвертом разделе мы описываем эмпирическую стратегию. После этого приведена интерпретация результатов. Наконец, в шестом разделе резюмируются результаты и формулируются ограничения проведенного исследования, а также приведены направления для дальнейших исследований.

Обзор литературы

Особенности авиарынков. Важной особенностью многих рынков авиаперевозок является так называемая система hub-and-spoke сетей: маршруты объединяют удаленные аэропорты с центральными аэропортами, хабами. В частности, если построить карту перелетов на данных Airfleets¹, то можно обнаружить, что российских авиакомпаний есть аэропорты-хабы в Москве и Санкт-Петербурге (рис. 1). Можно выделить две основные причины, по которым авиалинии выбирают такие сети. Во-первых, при увеличении присутствия авиалинии в аэропортах маршрута стоимость входа на маршрут падает, во-вторых, такие сети маршрутов мешают входу конкурентов (Aguirregabiria, Но, 2012). Кроме того, в России формированию альтернативной сети по принципу point-to-point препятствует недостаточный спрос на пассажирские перевозки в регионах (Скрылева, 2018).

Что касается влияния на потребителей, то эффект неоднозначный. С одной стороны, развитие сетей hub-and-spoke привело к большей частоте перелетов. С другой стороны, эти перелеты, как правило, менее удобные, так как потребителю необходимо сначала полететь в хаб, а затем в место назначения (Brueckner, Zhang, 2001; Berechman et al., 1998).

¹ Официальный сайт сервиса Airfleets. URL: <https://www.airfleets.net/home/> (дата обращения: 20.09.2023).

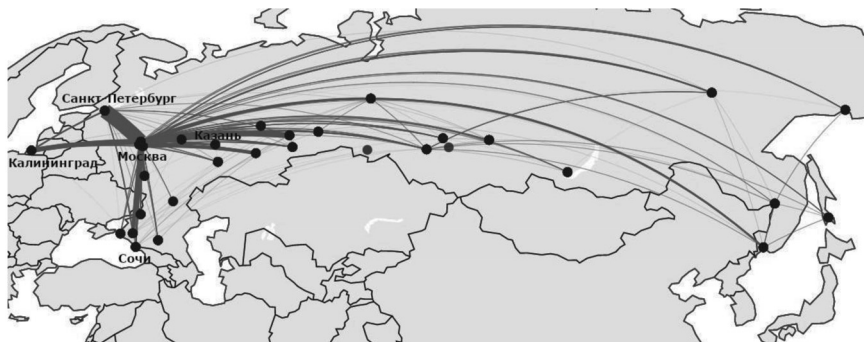


Рис. 1. Граф авиаперелетов внутри России с 2013 по 2018 г.

Источник: построено авторами.

Несмотря на то что схема hub-and-spoke потенциально может ухудшать условия конкуренции на рынках, вход новых участников и развитие систем бронирования (Starkie, 2012), а также совершенствование качества оказываемых услуг (Berry et al., 2019) приводят к усилению конкуренции на рынках пассажирских авиаперевозок. Рост конкуренции на маршруте может приводить к повышению благосостояния потребителей благодаря улучшению предоставляемых услуг. На американских данных показано, что компании, сталкивающиеся с ростом конкуренции, снижали долю отмененных рейсов и задержек рейсов (Gil, Kim, 2021). На данных об авиаперелетах в Бразилии показано, что снижение концентрации на рынке приводит к меньшему количеству задержек рейсов (Bendinelli et al., 2016).

Компании могут снижать цены также из-за конкурентного давления. В статье (Goolsbee, Syverson, 2008) показано, что компании значимо снижают тарифы на маршрутах, на которые с большой вероятностью может войти авиакомпания Southwest Airlines. Известно, что на авиарынках присутствуют ценовая дискриминация и разные цены в зависимости от меняющихся факторов издержек. Благодаря этим причинам авиарынкам свойственна большая дисперсия цен. Сначала считалось, что с ростом конкуренции дисперсия цен на авиарынках растет, однако позже при использовании несмещенных оценок было выявлено негативное влияние конкуренции на дисперсию цен (Borenstein, Rose, 1994; Gerardi, Shapiro, 2009).

Эффект от входа новых игроков на рынок зависит от того, насколько продукт нового игрока отличается от продуктов уже функционирующих компаний: чем больше похож продукт, тем сильнее эффект от конкуренции. Это видно, например, на примере рынка розничной торговли видеодисков, где участники рынка решением о географическом расположении «защищают себя от конкуренции» (Seim, 2006). Следовательно, повышение конкуренции мотивирует компании диверсифицировать свои продукты.

Кроме того, в зависимости от модели бизнеса авиакомпании могут изменять стратегии реакции на выход новых игроков (Маркова и др., 2019). В статье (Rupp, Liu, 2018) показано, что авиакомпании полного сервиса, конкурируя с лоукостерами, увеличивают расстояние между сидениями в самолете и увеличивают количество мест бизнес-класса.

С другой стороны, взаимодействие авиакомпаний на большом количестве рынков может упрощать их координацию и тем самым повышать вероятность сговора (Ciliberto, Williams, 2014).

Таким образом, исследования входа авиакомпаний на рынки обладают особенностями: с одной стороны, вход авиакомпаний на рынки зависит не только от структуры рынка, но и структуры сетей, которые построили существующие авиакомпании. С другой стороны, вход авиакомпаний на рынки неоднозначно влияет на конкуренцию, цены, качество перевозки на маршруте и даже вероятность сговора. При построении моделей входа компаний на авиарынки и интерпретации полученных результатов надо учитывать выделенные особенности. Перейдем к обзору структурной модели входа компаний на авиарынки.

Моделирование входа. Моделирование входа компаний на рынки в современной теории отраслевых рынков берет свое начало с работы (Hotelling, 1929), однако отправной точкой эмпирических моделей входа в теории отраслевых рынков считается статья (Bresnahan, Reiss, 1991). Базовая модель входа, представленная в этой статье, опирается на предположения о полной информации и гомогенности игроков. В этой модели мы увидим основные аспекты моделирования входа, а затем перейдем к обсуждению ее расширений.

Пусть есть n игроков, принимающих решение о входе на рынок: $a_i = \{0, 1\}$. Так как решения других игроков влияют на решение каждого игрока о входе на рынок, условие входа на рынок выглядит следующим образом:

$$\Pi^i(a_1, \dots, a_n) = \bar{\Pi}^i(a_1, \dots, a_n, X_i, \theta) - \epsilon^i(a_1, \dots, a_n) = \bar{\Pi}^i(X_i, \theta) - \epsilon_a^i \geq 0, \quad (1)$$

где $\bar{\Pi}^i$ — математическое ожидание прибыли фирмы i при условии входа на рынок, которая зависит от вектора наблюдаемых ковариат X_i , ненаблюдаемых параметров θ и n -мерного вектора a , отражающего действия всех игроков; ϵ — вектор случайных ошибок, отражающий все оставшиеся различия между издержками или функциями прибыли игроков. Игроки наблюдают вектор ϵ и обладают полной информацией о решениях друг друга. Обобщить функции платежей можно вектором $\Pi = \Pi(X, \theta) - \epsilon$.

Рассмотрим нахождение оптимального решения игрока. Стратегия a_i^* является наилучшим ответом на множество наилучших ответов всех игроков кроме a_{-i}^* , по Нэшу, если для $\forall a_i$ выполняется

$$\Pi^i(a_1^*, \dots, a_i^*, \dots, a_n^*) \geq \Pi^i(a_1^*, \dots, a_i, \dots, a_n^*). \quad (2)$$

Говоря неформально и применяя вышеприведенное определение к модели входа на рынок, оптимальным решением –й фирмы при заданных решениях о входе других фирм является вход, если прибыль от него больше, чем от воздержания от входа на рынок, и отказ от входа на рынок, если это принесет большую прибыль, чем вход на рынок.

Авиарынки относятся к рынкам с гетерогенными игроками, поэтому рассмотрение случая с гомогенными игроками является существенным ограничением модели (Bresnahan, Reiss, 1991). Тем не менее дальнейшие расширения наследуют из этой модели идею о существовании пороговых условий, при которых возможен вход новых компаний на рынок.

Исследования рынков с гетерогенными игроками, которые ослабляют предпосылку (Bresnahan, Reiss, 1991) об изолированных рынках, начинаются с работы (Berry, 1992). В идеальной модели в каждый период времени должен рассматриваться выбор всей сети возможных маршрутов каждой фирмой, но такое моделирование вряд ли представляется возможным, поэтому вместо общего равновесия в статье рассматривается частичное: в начале периода фирма принимает свою сеть маршрутов как заданную и принимает решение о входе на конкретный маршрут между двумя городами, а не рассматривает все возможные комбинации маршрутов. Рассмотрим модель, приведенную в данной статье.

Перечислим предпосылки, необходимые для существования равновесия Нэша в чистых стратегиях. Во-первых, прибыль фирмы должна убывать по количеству игроков на рынке, во-вторых, фирмы можно про-ранжировать по их прибыльности. При таких предпосылках равновесие существует, но оно не единственное, что приводит к эмпирическим трудностям, связанным с тем, что одинаковые условия могут приводить к разным результатам.

Распространенным подходом при моделировании ситуаций, в которых возникает несколько равновесий, является нахождение какого-то признака, общего для всех равновесий. Например, у всех равновесий в игре может быть одинаковое число фирм, выбравших войти на рынок, а сами множества таких фирм могут отличаться. Для уникальности числа фирм, которые вошли на рынок, необходимо наложить следующие ограничения на функции прибыли: гетерогенность фирм выражается только в фиксированной части прибыли, игра после входа является симметричной, прибыль убывает по количеству фирм на рынке. При такой функции прибыли все равновесия по Нэшу в чистых стратегиях имеют одинаковое количество фирм, принявших решение об участии в рынке.

Выводы теоретической модели С. Берри проверяет на данных об авиаперелетах в США, Origin and Destination Survey of Air Passenger Traffic¹.

¹ O&D survey // Bureau of Transportation Statistics. URL: <https://www.transtats.bts.gov/Homepage.asp> (дата обращения: 20.09.2023).

Эта выборка — примерно 10% от генеральной совокупности всех перелетов внутри США. Оценивается модель пробит, из которой делаются следующие выводы: население в городах маршрута положительно влияет на вероятность входа; зависимость от расстояния нелинейная (сначала вероятность растет, затем падает); вероятность входа на маршрут между туристическими городами при прочих равных условиях выше; переменные, характеризующие присутствие авиалинии в аэропортах маршрута, также значимо положительно влияют на вероятность входа на маршрут. Для оценки модели используется метод симулированного максимального правдоподобия. Знаки и значимость оценок коэффициентов не отличаются от оценок, полученных пробитом.

Такой же эмпирической стратегией и таким же набором переменных, но на европейских данных пользуются (Bontemps, Sampaio, 2020), результаты которых совпадают с результатами (Berry, 1992).

Предпосылка о независимости прибыльности фирмы от типов ее конкурентов является не очень реалистичной, особенно на рынке авиаперевозок, где как минимум можно выделить бюджетные авиакомпании и авиакомпании полного сервиса. Воздействие входа авиалинии на прибыли конкурентов может различаться в зависимости от авиалинии (гетерогенный эффект). В работе (Ciliberto, Tamer, 2009), как и в статье (Berry, 1992), моделируется игра с гетерогенными игроками и полной информацией, и на американских данных показано, что вход больших авиакомпаний (American, Delta, United) влияет на прибыли других компаний на маршруте не так, как вход бюджетных авиакомпаний (лоукостеров).

Еще один пример расширения модели входа, предложенной (Berry, 1992) — работа (Mazzeo, 2002). В модель входа добавляется эндогенная продуктовая дифференциация, и на примере рынка мотелей в США показано, что меньшее снижение прибыли от входа игрока с другим типом продукта дает объяснение дифференциации фирмами их продуктов. (Dunn, 2008) также включает выбор качества продукта в модель входа для анализа конкуренции между авиалиниями, предлагающими перелет без пересадок и перелет с пересадкой в аэропорте-хабе. Автор показывает, что присутствует эффект каннибализации: авиалиния, предлагающая перелет с пересадкой, с меньшей вероятностью предложит на этом маршруте перелет без пересадки. Кроме того, очевидно, есть эффект конкуренции: вход конкурента, предлагающего перелет без пересадок на маршруте, снижает вероятность авиалинии предложить на этом маршруте перелет без пересадок.

Моделирование входа фирмы на рынок может быть не только конечной задачей исследования, но и его составной частью: например, решения о входе и последующие решения о ценообразовании могут быть частями одной модели. В статье (Ciliberto et al., 2021) показано, что включение в модель эндогенного входа фирм на рынок приводит к более корректной оценке эластичности спроса.

В обзоре литературы были приведены особенности конкуренции на рынках авиаперевозок, базовая модель входа, ее расширения и результаты эмпирических оценок моделей входа на преимущественно американских и европейских данных. В следующем разделе будут приведены методики сбора и анализ российских данных.

Данные и разведочный анализ

Для оценки модели входа на рынки пассажирских авиаперевозок был собран авторский датасет с данными о перелетах, характеристиках рынков и присутствия авиакомпаний на рынках. В отличие от, например, США, исторические данные об авиаперелётах внутри России найти в открытом доступе достаточно сложно¹. Для сбора данных об авиаперелетах использовалась информация с сайта авиалюбителей^{2, 3}. На рис. 2 для примера показана часть результатов по запросу «Домодедово — Пулково». Собранные на сайте данные представляют собой маленькую и, вероятно, смещенную выборку, поэтому выводы дальнейшего исследования стоит очень аккуратно экстраполировать на генеральную совокупность. Однако более объемных и репрезентативных данных по России в открытом доступе не найдено.

Так, было получено 4526 наблюдений об авиакомпаниях, дате и маршруте перелета с 1988 по 2022 г., больше всего наблюдений с 2013 по 2021 г. (рис. 3). Для оценки моделей будут использоваться данные с 2013 по 2018 г., так как за эти периоды больше всего данных, и они были до пандемии, которая повлияла на решения фирм и вызвала структурные сдвиги спроса. Выбор такого периода, с одной стороны, позволяет сравнить полученные результаты с результатами предшествующих исследований (в которых используются данные, не подверженные значительному влиянию внешних шоков). С другой стороны, выбор данных за более ранние периоды позволяет на этом этапе говорить о выполнении предпосылки о наличии равновесия на рынках во всех исследуемых периодах (предпосылки, которая не выполняется для пандемийных лет, а также для санкционных периодов после февраля 2022 г.⁴).

¹ Степень агрегации данных из альтернативных источников информации, которые находятся в открытом доступе (в частности, данные ФАВТ о количестве перевезенных пассажиров в разрезе авиакомпаний и данные о количестве перевезенных пассажиров в аэропорту), не позволяет их использовать для целей данной статьи.

² Официальный сайт сервиса Airfleets. URL: <https://www.airfleets.net/home/> (дата обращения: 20.09.2023).

³ Сайт защищен от автоматического парсинга данных системой CloudFlare, поэтому все файлы формата HTML были скачаны на компьютер, затем мы парсили таблицы с авиаперелетами из файлов HTML.

⁴ В дальнейших исследованиях полученные для ранних периодов данные могут быть использованы для сравнения с потенциально неравновесными периодами.

Flightlog Moscow Domodedovo St Petersburg Pulkovo

102 flights found for your search

Airline : all
Departure airport : Moscow Domodedovo Russia

Date : all

Aircraft : all
Arrival airport : St Petersburg Pulkovo Russia

Page 1/6 [Next page]

1 2 3 4 5 6
[Search]

Date	Flight	Departure	Arrival	Aircraft	Recorded by
25/03/2022	View on the map S7 1013	DME Moscow Domodedovo - Russia	LED St Petersburg Pulkovo - Russia	VQ-BVV E170 MSN 48	AlIAirliners Public profile of AlIAirliners
26/02/2022	View on the map S7 1013	DME Moscow Domodedovo - Russia	LED St Petersburg Pulkovo - Russia	VQ-BRK B737NG MSN 41708	barsey Public profile of barsey
01/12/2021	View on the map S7 1015	DME Moscow Domodedovo - Russia	LED St Petersburg Pulkovo - Russia	VP-BLD B737NG MSN 35283	avk111 Public profile of avk111
30/05/2021	View on the map S7 1025	DME Moscow Domodedovo - Russia	LED St Petersburg Pulkovo - Russia	VP-BOL A320 MSN 6066	syntez Public profile of syntez
27/03/2021	View on the map S7 1013	DME Moscow Domodedovo - Russia	LED St Petersburg Pulkovo - Russia	VQ-BRQ B737NG MSN 41710	barsey Public profile of barsey
27/02/2021	View on the map S7 1013	DME Moscow Domodedovo - Russia	LED St Petersburg Pulkovo - Russia	VP-BWM A320 MSN 8852	barsey Public profile of barsey
19/12/2020	View on the map S7 1011	DME Moscow Domodedovo - Russia	LED St Petersburg Pulkovo - Russia	VQ-BET A320 MSN 4150	Ashor Public profile of Ashor

Рис. 2. Пример выдачи сайта Airfleets

Источник: построено авторами.

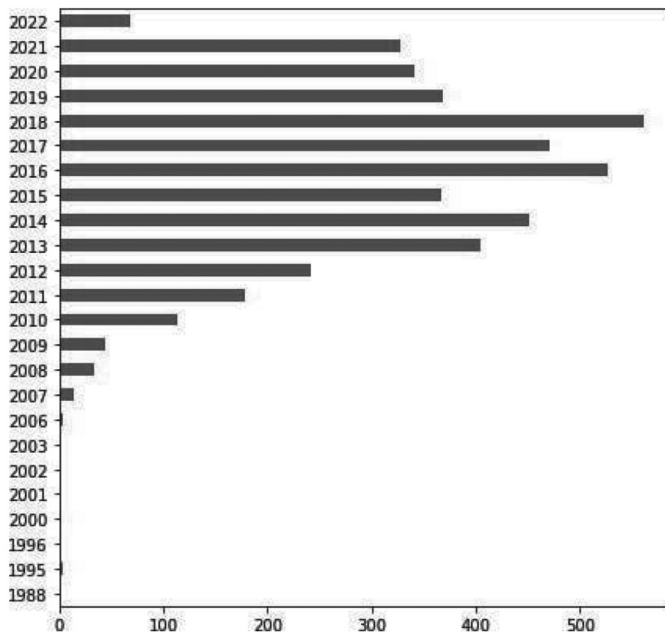


Рис. 3. Распределение наблюдений по годам

Источник: построено авторами.

В датасете присутствуют наблюдения о вылетах из 29 уникальных аэропортов и прилетах в 37 уникальных аэропортов, причем первое множество является подмножеством второго: всего 37 уникальных аэропортов в данных. Фактически в России 91 аэропорт федерального назначения¹, соответственно, в выборку попало примерно 40% аэропортов, и это преимущественно большие аэропорты.

Кроме данных об авиаперелётах для исследования необходимы данные, характеризующие рынки, — маршруты между двумя городами. В литературе в качестве факторов входа используются следующие характеристики рынков: население городов, в которых расположены аэропорты, расстояние между аэропортами и переменная, характеризующая, являются ли города туристическими. Исходя из предпосылки, что за период с 2013 по 2018 г. соотношение населения городов России сильно не менялось, используются данные о населении городов на 1 января 2018 г. с сайта, использующего данные службы Росстата². Для нахождения географических координат аэропортов использовались сайты *prokerala* и *ourairports*^{3, 4}. Затем между координатами рассчитывалось геодезическое расстояние — кратчайшее расстояние между двумя точками на поверхности эллипсоидной модели Земли. Для определения туристических городов России использовалось два источника: рейтинг туристических городов России, созданный на основе поисковых запросов «достопримечательности...» и «что посмотреть в...» в Яндексe за 2018 и 2019 гг., и данные Росстата о туристической индустрии в России^{5, 6}. На основании обоих источников выбраны Москва, Санкт-Петербург, Сочи и Казань, так как эти города оказались самыми туристическими с большим отрывом от остальных.

Наряду с общепринятыми переменными в данном исследовании предложены следующие переменные, характеризующие маршрут: найденные на сайтах *prokerala* и *ourairports* типы аэропортов (Middle или Large) и являются ли города, в которых расположены аэропорты, миллионниками.

В качестве переменных, характеризующих авиакомпанию на рынке, взяты количество аэропортов, которые авиакомпания обслуживала в паре

¹ Перечень аэропортов федерального значения // ФАВТ. URL: <https://favt.gov.ru/deyatelnost-ajerorporty-i-ajerodromy-perechen-aeroporov/> (дата обращения: 21.09.2023).

² Города России // Сайт Города-Россия. URL: <https://города-россия.рф> (дата обращения: 21.09.2023).

³ Russia Airports & List of Airports in Russia — Russia Airport Codes // *prokerala*. URL: <https://www.prokerala.com/travel/airports/russia/> (дата обращения: 21.09.2023).

⁴ OurAirports // *ourairports*. URL: <https://ourairports.com/> (дата обращения: 21.09.2023).

⁵ Самые популярные города России // Портал TownTravel. URL: <http://towntravel.ru/interesnie-fakti-o-gorodah-rossyi/samyie-populyarnye-goroda-rossii-top-50-za-2018-2019-god.html> (дата обращения: 21.09.2023).

⁶ Данные Росстата о туризме // Официальный сайт Росстата. URL: <https://rosstat.gov.ru/folder/313/document/100185> (дата обращения: 21.09.2023).

с аэропортами маршрута в прошлом периоде, и количество всех аэропортов, которые авиакомпания обслуживала в прошлом периоде. Приведем пример. Пусть авиакомпания обслуживает в 2017 г. маршруты АВ, ВС, DC, EA, FG, тогда для маршрута AC в 2018 г. первая переменная будет равна 4 (аэропорты В, Е для аэропорта А, аэропорты В, D для аэропорта С), а вторая — 7 (А, В, С, D, Е, F, G) (рис. 4).

Для создания датасета о входе фирм на маршруты использовалась следующая стратегия. В одной модели используется два соседних года, так как за этот период не должны сильно измениться факторы спроса и предложения, при этом фирма успевает за такой временной промежуток принять решение о входе на рынок¹. Рассмотрим для примера периоды с 2017 по 2018 г. Как уже было отмечено выше, рынок — это ненаправленный маршрут между двумя городами. Направление маршрутов не имеет значения в нашей модели, поэтому мы приравниваем маршруты АВ и ВА к АВ путем сортировки. Декартово произведение всех аэропортов, которые встречались в маршрутах 2017–2018 гг., и всех авиакомпаний, которые встречались в наблюдениях 2018 г., дает нам все возможные маршруты для всех компаний, обслуживающих аэропорты в 2018 г.

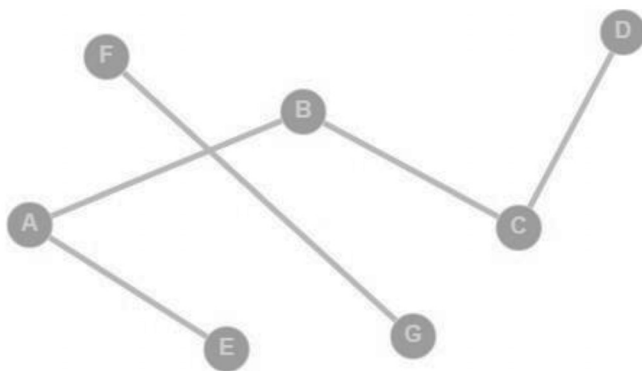


Рис. 4. Пример маршрутов для иллюстрации переменных, характеризующих присутствие авиалинии на рынках

Источник: построено авторами.

Объясняемой переменной является индикатор входа. Если авиакомпания присутствует на маршруте в 2018 г., но не была на нем в 2017 г., то в 2018 г. авиакомпания вошла на маршрут. Подобные датасеты построены для периодов 2013–2014, 2014–2015, 2015–2016, 2016–2017 и 2017–2018 гг.

¹ В авиаперелетах есть сезонность, но для двух соседних годов ее влияние исключается, поэтому в качестве единицы наблюдения выбрано наблюдение о присутствии авиакомпании на маршруте в этом году.

Для описания выборки и проверки ошибок в данных приведена описательная характеристика данных (табл. 1).

Так как вход авиакомпании на маршрут — редкое событие в полученных нами данных, во всех используемых наборах данных наблюдается ярко выраженный дисбаланс классов: доля входов составляет 0,6%. На этапе разведочного анализа данных мы проанализируем, как выбранные переменные помогают разграничить значения объясняемой переменной.

Таблица 1

Описательные статистики объясняющих переменных

	count	mean	std	min	50%	Max
served_airports_pr_period (количество аэропортов в t-1)	51 894	4,68	6,28	0	2	26
neigh_airports_pr_period (количество соседних аэропортов в t-1)	51 894	0,53	1,7	0	0	26
airport_type (количество больших аэропортов на маршруте)	51 894	0,59	0,63	0	1	2
population (прокси потенциального объема спроса на маршруте)	51894	1234,23	1186,18	75,5	838,92	8181,2
megapolis (количество городов-миллионников на маршруте)	51 894	1,06	0,69	0	1	2
tourist (количество туристических городов на маршруте)	51 894	0,38	0,54	0	0	2
distance_km (расстояние между аэропортами)	51 894	2774,33	2008,95	135,75	2087,11	7492,86

Источник: построено авторами.

Среднее геометрическое населения в двух городах маршрута является переменной, характеризующей спрос. На рис. 5 показано, что маршруты, на которые входят авиакомпании, расположены между более густонаселенными городами.

Эта закономерность видна и при построении таблицы сопряженности переменной, равной числу городов-миллионников на маршруте, и индикатора входа: авиакомпании предпочитают маршруты между городами-миллионниками (табл. 2). Количество туристических городов на маршруте еще лучше разграничивает классы (показывает различия между маршрутами, на которые входят авиалинии, и другими маршрутами): авиалинии чаще входят на маршруты, на которых есть туристический город (табл. 3).

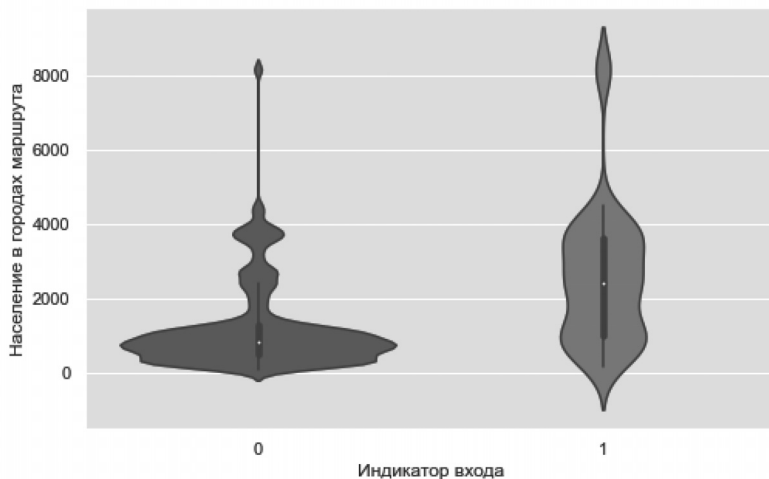


Рис. 5. Распределение населения в зависимости от того, был (справа) или не был (слева) вход на рынок
Источник: построено авторами.

Таблица 2

Таблица сопряженности для переменных входа и количества мегаполисов на маршруте

<i>n</i> городов-миллионников / индикатор входа	0	1	2
0	0,21	0,52	0,27
1	0,08	0,48	0,44

Источник: построено авторами.

Таблица 3

Таблица сопряженности для переменных входа и количества туристических городов на маршруте

<i>n</i> туристических городов / индикатор входа	0	1	2
0	0,65	0,33	0,03
1	0,22	0,61	0,17

Источник: построено авторами.

Расстояние между аэропортами влияет и на спрос, и на предложение. Фактором спроса расстояние является ввиду того, что чем больше расстояние между городами, тем больше спрос на авиаперевозки относительно других видов транспорта: поездов и автомобилей. При этом расстояние влияет на количество потребляемого топлива, следовательно, является фактором издержек. Исследования на американских и европейских данных показывали нелинейное влияние расстояния на вероятность входа (парабола ветвями вниз). Однако география России отличается от Европы и США, протяженность России с запада на восток составляет более 10 тыс. км. На многие протяженные маршруты авиалинии без пересадок не входят, например, от Москвы до Петропавловска-Камчатского без пересадок долететь нельзя. Соответственно, вероятность входа должна падать с ростом расстояния. На маршрутах, где входа не было, распределение расстояния более равномерное, чем на маршрутах, где вход был (рис. 6). Во втором случае мода распределения равна примерно 1800 км.

Еще одной переменной, характеризующей и спрос, и предложение, является тип аэропорта (большой или средний). Как видно из табл. 4, распределение типов аэропортов отличается в зависимости от того, был ли вход на маршрут: авиакомпании чаще входят на маршруты между большими аэропортами. Это можно объяснить и спросом, и предложением: с одной стороны, вероятно, большие аэропорты предоставляют больше места для стоянки самолетов, с другой — большие аэропорты расположены в более крупных городах с большим населением.

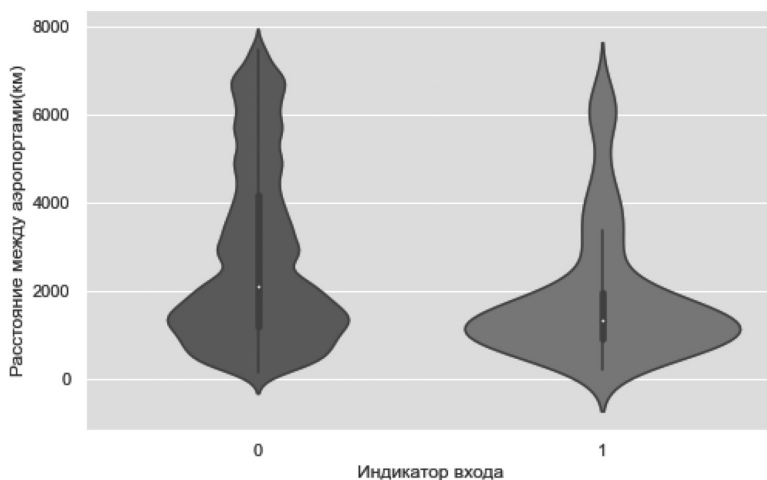


Рис. 6. Распределение расстояния в зависимости от того, был (справа) или не был (слева) вход на рынок

Источник: построено авторами.

**Таблица сопряженности для переменных входа и типов аэропортов
на маршруте**

<i>n</i> больших аэропортов / индикатор входа	0	1	2
0	0,49	0,43	0,08
1	0,25	0,54	0,21

Источник: построено авторами.

Исходя из обзора литературы и предварительного анализа данных, можно выдвинуть гипотезы о предполагаемом влиянии переменных на вероятность входа. Количество аэропортов, обслуживаемых авиакомпаниями в паре с аэропортами маршрута в прошлом периоде, и количество аэропортов, обслуживаемых авиакомпаниями в прошлом периоде — метрики присутствия авиакомпании на конкретном рынке и на всех рынках в целом соответственно. Ожидается, что эти переменные будут положительно влиять на вероятность входа. Прокси-переменные спроса (население, количество туристических городов на маршруте, количество городов-миллионников, количество больших аэропортов) характеризуют рынки, более привлекательные для входа авиакомпаний, соответственно, должны положительно влиять на объясняемую переменную. Наконец, расстояние между аэропортами в исследованиях на американских и европейских данных нелинейно (парабола ветвями вниз) влияет на вероятность входа, поэтому ожидаем нелинейное влияние в нашем исследовании.

Эмпирическая стратегия

Для проверки сформулированных в предыдущем разделе гипотез используется описанную в обзоре литературы модель (Berry, 1992)¹, для которой были выбраны следующие переменные, характеризующие рынки и деятельность фирмы на рынках в предыдущем периоде:

- `neigh_airports_pr_period` — количество аэропортов, обслуживаемых авиакомпаниями в паре с аэропортами маршрута в прошлом периоде;

¹ Выбор классической модели входа на авиарынки в качестве основы для эмпирической части позволит избежать рисков, связанных с ошибками в оценках из-за особенностей используемых данных: такой выбор позволяет сравнить полученные результаты с большим количеством предшествующих исследований, что позволит говорить о надежности полученных результатов.

- `served_airports_pr_period` — количество аэропортов, обслуживаемых авиакомпаниями в прошлом периоде;
- `airport_type` — количество больших аэропортов на маршруте;
- `population` — среднее геометрическое населения в городах маршрута в тысячах человек (прокси потенциального объема спроса на маршруте);
- `megapolis` — количество городов-миллионников на маршруте;
- `tourist` — количество туристических городов на маршруте;
- `distance_km` — геодезическое расстояние между аэропортами в километрах.

Предполагается, что расстояние нелинейно влияет на вероятность входа на маршрут, поэтому кроме предложенных выше переменных в исследованиях традиционно используется квадрат расстояния. Однако на наших данных оценка коэффициента при квадрате расстояния была незначима, и при добавлении этой переменной влияние переменной расстояния тоже становилось незначимым. Кроме того, при добавлении квадрата расстояния появляется мультиколлинеарность (фактор инфляции дисперсии (VIF) у переменных расстояния и квадрата расстояния равен 17), которой иначе нет.

Для выявления факторов, влияющих на вероятность входа фирмы на рынок, мы решаем задачу бинарного выбора. Для оценки точности прогнозов полученных моделей вся выборка делится на обучающую и тестовую выборки в соотношении 3 : 1 случайным образом, однако соотношение значений объясняемой переменной в обеих выборках одинаковое. На обучающей выборке все модели обучаются, а на тестовых данных мы смотрим значения метрик бинарной классификации. Для оценки адекватности прогнозов выбраны следующие метрики:

- `accuracy` — доля правильных ответов алгоритма;
- `precision` (точность) — доля объектов какого-то класса среди объектов, для которых модель спрогнозировала этот класс;
- `recall` (полнота) — доля объектов, для которых модель спрогнозировала какой-то класс, среди всех объектов этого класса;
- `f1-measure` — среднее гармоническое точности и полноты.

Из эконометрических методов оценивания моделей входа традиционно используется пробит, с которого начинается эмпирическая стратегия данной работы (Berry, 1992; Bontemps, Sampaio, 2020). Кроме того, мы используем алгоритмы машинного обучения, которые могут выявить нелинейное влияние переменных интереса на зависимую и дают альтернативные способы оценки важности переменных, что позволит говорить об устойчивости полученных результатов. Однако в данных очень маленькая доля положительного класса (наблюдений, где вход был), и, как следствие, модели могут не выявить закономерностей в данных, объясняющих влияние переменных интереса на вероятность входа, так как вход фирмы

на рынок — редкое событие¹. Поэтому используем четыре метода исправления дисбаланса классов, после каждого из которых оцениваем три модели машинного обучения: логит, случайный лес и градиентный бустинг.

Первый метод исправления дисбаланса классов заключается в следующем. При построении модели машинного обучения подбираются параметры, минимизирующие функционал ошибки. При ярко выраженном дисбалансе классов, который присутствует в данных, на подбор параметров гораздо сильнее влияют наблюдения, в которых входа не было, так как они — подавляющее большинство выборки. Для того чтобы функционал ошибки в равной мере учитывал оба класса, используются веса, обратно пропорциональные частотности классов.

Другие три метода работы с данными при дисбалансе классов — это разные способы пересемплирования (*Oversampling the minority class*). После использования этих методов количество наблюдений, в которых входа было, становится равным количеству наблюдений, в которых входа не было. Первым таким способом является случайное пересемплирование, при котором мы создаем новые наблюдения меньшего класса, случайно семплируя объекты меньшего класса с возвращением. Так, благодаря дубликатам первого класса соотношение классов в выборке становится равным.

Кроме метода случайного пересемплирования есть еще два популярных метода: *Synthetic Minority Oversampling Technique (SMOTE)* и *Adaptive Synthetic sampling method (ADASYN)* (Chawla et al., 2002; He et al., 2008). Метод SMOTE работает следующим образом: на отрезках между наблюдением из малого класса и каждым из его k ближайших соседей (тоже принадлежащих малому классу) случайным образом создаются новые наблюдения малого класса. Идея метода ADASYN аналогична методу SMOTE, но с одним изменением: количество новых точек, которые создаются с помощью какого-то наблюдения малого класса, пропорционально числу объектов другого класса в окрестности этого наблюдения.

Все вышеперечисленные модели обучаются в сочетании с разными способами исправления дисбаланса классов, затем их метрики на тестовых данных используются для оценки адекватности моделей.

Результаты

Из-за ярко выраженного дисбаланса классов большинство прогнозов всех моделей — нули (отсутствие входа), поэтому на тестовых данных у моделей высокая доля правильных ответов². Однако усредненные

¹ Проблема с оценками редких событий заключается в переоценке вероятности отсутствия входа — с похожей проблемой сталкиваются и другие исследователи, использующие данные с редкими событиями (например, банкротствами (Донец, Могилат, 2017)).

² Объясняемая переменная в подавляющем большинстве наблюдений равна нулю, также почти все прогнозы равны нулю, поэтому большинство прогнозов верны.

по классам метрики recall, precision, f1-score принимают маленькие значения из-за низких полноты и точности на данных положительного класса (наличие входа). Метрики моделей на тестовых данных принципиально не отличаются, поэтому нельзя говорить о превосходстве одних моделей над другими. Все метрики приведены в приложении А.

Далее мы интерпретируем коэффициенты пробит-модели (табл. 5) и сравниваем их с результатами, полученными при построении моделей машинного обучения (Приложение Б).

Исходя из оценок пробит-модели, количество аэропортов, обслуживаемых авиакомпанией в прошлом периоде, не влияет значимо на решение авиакомпании о входе на маршрут. Модели машинного обучения также не выявили заметного влияния этой переменной на вероятность входа. Количество аэропортов, обслуживаемых авиакомпанией в паре с аэропортами маршрута в прошлом периоде, значимо и положительно влияет на вероятность входа. Построение моделей машинного обучения подтверждает сильную взаимосвязь этой переменной с вероятностью входа. Чем больше узлы маршрута похожи на хабы авиалинии, тем выше вероятность появления маршрута между этими узлами в следующем периоде. Количество больших аэропортов на маршруте значимо и положительно влияет на вероятность входа, однако при построении моделей машинного обучения эта переменная не вносит особого вклада в прогнозы. Скорее всего, это может быть связано с тем, что большинство крупных аэропортов уже используются авиакомпаниями как хабы.

Таблица 5

Коэффициенты пробит-модели

	Коэффициент	Стандартная ошибка
константа	-2,7386***	0,085
served_airports_pr_period (количество аэропортов в t-1)	0,0028	0,004
neigh_airports_pr_period (количество соседних аэропортов в t-1)	0,0839***	0,007
airport_type (количество больших аэропортов на маршруте)	0,0883*	0,046
population (прокси потенциального объема спроса на маршруте)	0,0001***	0
megapolis (количество городов-миллионников на маршруте)	-0,1014**	0,049
tourist (количество туристических городов на маршруте)	0,3124***	0,057
distance_km (расстояние между аэропортами)	-0,0001***	0

Примечание: зависимая переменная — индикатор входа. R-квадрат МакФаддена = 0,16. n = 38 920. * p-значение < 0,1. ** p-значение < 0,05. *** p-значение < 0,01.

Источник: построено авторами.

Среднее геометрическое населения в городах маршрута (population), которое можно считать одной из прокси-переменных спроса, значимо и положительно влияет на вероятность входа на маршрут, что согласуется с теорией и предыдущими исследованиями. Оценивание логита, случайного леса и градиентного бустинга подтверждает эту взаимосвязь.

Переменная, равная количеству городов-миллионников на маршруте (megapolis), значимо отрицательно влияет на вероятность входа. Это может быть связано с тем, что наиболее «удобные» маршруты между городами-миллионниками уже являются активными для авиакомпаний, а вход на новые может быть связан со снижением пассажиропотока: например, между Уфой и Воронежем нет прямых рейсов, а перелеты осуществляются через Москву. Это связано с тем, что добавление нового маршрута не через Москву может снизить загрузку на всех маршрутах, связывающих эти города. Коэффициенты логит-моделей при этой переменной также отрицательны, а при построении градиентного бустинга и случайного леса у этой переменной маленькая важность.

Количество туристических городов на маршруте (tourist) — хороший индикатор спроса, поэтому большее количество туристических городов повышает вероятность входа авиакомпаний на маршрут, что согласуется с другими исследованиями.

Ввиду особенностей географии России, а именно большой протяженности с Запада на Восток, результат оценки влияния расстояния между аэропортами (distance_km) отличается от результатов исследований на американских и европейских данных. Полученные оценки говорят о значимом отрицательном влиянии расстояния на вероятность входа на маршрут. Как было отмечено ранее, на многие протяженные маршруты внутри России авиалинии без пересадок не входят, например, от Москвы до Петропавловска-Камчатского без пересадок долететь нельзя. Соответственно, падение вероятности входа на маршрут с ростом его протяженности является ожидаемым результатом на российских данных.

Обсуждение полученных результатов

Полученные оценки влияния характеристик маршрутов на вероятность входа могут быть полезны при анализе конкуренции на рынках и разработки политики в области стимулирования авиакомпаний входить на новые маршруты. Судя по результатам исследования, авиакомпания предпочитает входить на не очень протяженные маршруты между туристическими, густонаселенными городами, в аэропортах которых авиакомпания уже присутствует. Этот результат может быть использован для согласования действий политики в области туризма и развития пассажирских перевозок: например, в 2022 г. популярным направлением стали курорты Кавказа, поэтому часть финансирования полетов может быть перенесена

на более важные и менее привлекательные для авиакомпаний маршруты. Кроме того, из-за недостаточных объемов спроса авиакомпании могут принимать решение о входе на менее протяженные маршруты, что важно в текущих условиях, так как позволяет корректировать запросы на создание и производство самолетов с разной дальностью полета: потенциально в условиях санкций и с учетом полученных в данном исследовании результатов можно говорить о высокой потребности в среднемагистральных самолетах.

Важно отметить, что полученные результаты следует интерпретировать осторожно, так как в исследовании присутствует ряд ограничений. Так, из-за ограничений на данные ряд факторов, которые могут влиять на вероятность входа, не были учтены. Например, финансовые характеристики авиакомпаний, индикатор наличия скоростных поездов между городами маршрута, пассажирооборот компании в аэропортах маршрута. Собранная выборка содержит данные о перелетах преимущественно между большими аэропортами, что не позволяет делать вывод о вероятности входа авиакомпаний на маршруты, где одной из точек является малый аэропорт. Из-за того, что вход на маршрут является редким событием, модели обучались в основном на данных, где входа не было. Пересемплирование выборки позволило частично решить эту проблему.

Заключение

В данной работе выявлены факторы, влияющие на решение фирмы о входе на новый рынок пассажирских авиаперевозок (ненаправленный маршрут между двумя аэропортами) внутри России. Показано, что наибольший положительный вклад вносят количество аэропортов, обслуживаемых авиалинией в паре с аэропортами маршрута в прошлом периоде, население, количество туристических городов. Отрицательный вклад вносит расстояние между аэропортами. Полученные результаты частично согласуются с результатами исследований, проведенных на европейских и американских данных. Это позволяет говорить о надежности полученных результатов. При этом результат относительно отрицательного вклада фактора расстояния между аэропортами на вероятность входа авиакомпании позволяет обсуждать особенности российской отрасли авиаперевозок и формулировать рекомендации для ее поддержки.

Можно выделить следующие направления для дальнейших исследований. Во-первых, для более достоверных оценок стоит собрать более полные данные о перелетах внутри России. Во-вторых, используемый в данной работе набор данных может быть использован для оценки с помощью симуляций, что не будет требовать предпосылки о независимости прибыли фирмы от числа конкурентов. В-третьих, имеет смысл добавить другие переменные, характеризующие фирмы и рынки: наличие скорост-

ных поездов между городами маршрута, принадлежность авиакомпаний государству или альянсам. В-четвертых, на российских данных об авиарынках ранее не оценивались динамические модели, и это тоже может стать одним из направлений дальнейших исследований.

Список литературы

Донец, С. А., & Могилат, А. Н. (2017). Кредитование и финансовая устойчивость российских промышленных компаний: микроэкономические аспекты анализа. *Деньги и кредит*, (7), 41–51.

Комаристый, Е. Н. (2004). Математические подходы к анализу спроса на пассажирские авиаперевозки. *Маркетинг и маркетинговые исследования*, (3), 10–16.

Лукиянов, С. А., Тиссен, Е. В., & Кисляк, Н. (2007). Рынок авиационных пассажирских перевозок России: квазиконкуренция или.... *Вопросы экономики*, 11, 120–138. <https://doi.org/10.32609/0042-8736-2007-11-120-138>

Маркова, О. А., Мелешкина, А. И., & Моросанова, А. А. (2019). Особенности конкурентных стратегий авиакомпаний с учетом тарифообразования. *Научные исследования экономического факультета. Электронный журнал*, 11(4 (34)), 19–34.

Матюха, С. В. (2022). Развитие модели регионального рынка авиаперевозок. Потребители рынка. *Транспортное дело России*, (4).

Скрылева, Е. В. (2018). Исследование факторов, влияющих на развитие региональных авиаперевозок. *Инновации и инвестиции*, (9), 236–240.

Aguirregabiria, V. (2012). Empirical industrial organization: models, methods, and applications. *University of Toronto, Preliminary version*.

Aguirregabiria, V., & Ho, C. Y. (2012). A dynamic oligopoly game of the US airline industry: Estimation and policy experiments. *Journal of Econometrics*, 168(1), 156–173. <https://doi.org/10.1016/j.jeconom.2011.09.013>

Bindinelli, W. E., Bettini, H. F., & Oliveira, A. V. (2016). Airline delays, congestion internalization and non-price spillover effects of low cost carrier entry. *Transportation Research Part A: Policy and Practice*, 85, 39–52. <https://doi.org/10.1016/j.tra.2016.01.001>

Berechman, J., Poddar, S., & Shy, O. (1998). Network structure and entry in the deregulated airline industry. *Keio economic studies*, 35(2), 71–82.

Berry, S. T. (1992). Estimation of a Model of Entry in the Airline Industry. *Econometrica: Journal of the Econometric Society*, 889–917.

Berry, S., Gaynor, M., & Scott Morton, F. (2019). Do increasing markups matter? Lessons from empirical industrial organization. *Journal of Economic Perspectives*, 33(3), 44–68. <https://doi.org/10.1257/jep.33.3.44>

Berry, S., & Reiss, P. (2007). Empirical models of entry and market structure. *Handbook of industrial organization*, 3, 1845–1886.

Bontemps, C., & Sampaio, R. M. (2020). Entry games for the airline industry 1. *Air Transport and Regional Development Methodologies*, 226–248.

Borenstein, S., & Rose, N. L. (1994). Competition and price dispersion in the US airline industry. *Journal of Political Economy*, 102(4), 653–683. <https://doi.org/10.1086/261950>

Bresnahan, T. F., & Reiss, P. C. (1991). Empirical models of discrete games. *Journal of Econometrics*, 48(1-2), 57–81. [https://doi.org/10.1016/0304-4076\(91\)90032-9](https://doi.org/10.1016/0304-4076(91)90032-9)

Brueckner, J. K., & Zhang, Y. (2001). A model of scheduling in airline networks: how a hub-and-spoke system affects flight frequency, fares and welfare. *Journal of Transport Economics and Policy (JTEP)*, 35(2), 195–222.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16, 321–357. <https://doi.org/10.1613/jair.953>

Ciliberto, F., Murry, C., & Tamer, E. (2021). Market structure and competition in airline markets. *Journal of Political Economy*, 129(11), 2995–3038.

Ciliberto, F., & Tamer, E. (2009). Market structure and multiple equilibria in airline markets. *Econometrica*, 77(6), 1791–1828. <https://doi.org/10.3982/ECTA5368>

Ciliberto, F., & Williams, J. W. (2014). Does multimarket contact facilitate tacit collusion? Inference on conduct parameters in the airline industry. *The RAND Journal of Economics*, 45(4), 764–791. <https://doi.org/10.1111/1756-2171.12070>

Cleeren, K., Verboven, F., Dekimpe, M. G., & Gielens, K. (2010). Intra- and interformat competition among discounters and supermarkets. *Marketing science*, 29(3), 456–473. <http://dx.doi.org/10.1287/mksc.1090.0529>

Dunn, A. (2008). Do low-quality products affect high-quality entry? Multiproduct firms and nonstop entry in airline markets. *International Journal of Industrial Organization*, 26(5), 1074–1089. <https://doi.org/10.1016/j.ijindorg.2007.10.001>

Einav, L., & Levin, J. (2010). Empirical industrial organization: A progress report. *Journal of Economic Perspectives*, 24(2), 145–162.

Gerardi, K. S., & Shapiro, A. H. (2009). Does competition reduce price dispersion? New evidence from the airline industry. *Journal of Political Economy*, 117(1), 1–37. <http://dx.doi.org/10.1086/597328>

Gil, R., & Kim, M. (2021). Does competition increase quality? Evidence from the US airline industry. *International Journal of Industrial Organization*, 77, 102742. <https://doi.org/10.1016/j.ijindorg.2021.102742>

Goolsbee, A., & Syverson, C. (2008). How do incumbents respond to the threat of entry? Evidence from the major airlines. *The Quarterly journal of economics*, 123(4), 1611–1633. <https://doi.org/10.1162/qjec.2008.123.4.1611>

Grieco, P. L. (2014). Discrete games with flexible information structures: An application to local grocery markets. *The RAND Journal of Economics*, 45(2), 303–340. <https://doi.org/10.1111/1756-2171.12052>

He, H., Bai, Y., Garcia, E. A., & Li, S. (2008, June). ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)* (p. 1322–1328). IEEE.

Hotelling, H. (1990). Stability in competition. In *The collected economics articles of Harold Hotelling* (p. 50–63). Springer, New York, NY.

Lukyanov, S. A., Ruzhanskaya, L. S., Avramenko, E. S., & Stroeve, V. V. (2018). Restraints on competition in the Russian air passenger market. *St Petersburg University Journal of Economic Studies*, 34(1), 134–148. <https://doi.org/10.21638/11701/spbu05.2018.107>

Lukyanov, S., Tissen, E., & Kislyak, N. (2009). The Russian Airline Industry: Contestable Market Or...? *Perspectives of Innovations, Economics, and Business*, 2(1231-2016-100855), 30–33.

Mazzeo, M. J. (2002). Product choice and oligopoly market structure. *RAND Journal of Economics*, 221–242.

Rupp, N. G., & Liu, A. N. (2018). Product quality choices and competition. *Journal of Transport Economics and Policy (JTEP)*, 52(4), 367–390.

Seim, K. (2006). An empirical model of firm entry with endogenous product-type choices. *The RAND Journal of Economics*, 37(3), 619–640. <https://doi.org/10.1111/j.1756-2171.2006.tb00034.x>

Starkie, D. (2012). European airports and airlines: Evolving relationships and the regulatory implications. *Journal of Air Transport Management*, 21, 40–49. <https://doi.org/10.1016/j.jairtraman.2011.12.016>

References

Donets, S., & Mogilat, A. (2017). Lending to russian industrial companies and their financial soundness: microeconomic aspects of the study. *Russian Journal of Money & Finance*, 7, 41–51.

Komaristiy, E. N. (2004). Mathematical approaches to the analysis of demand for passenger air transportation. *Marketing and marketing research*, 3, 10–16.

Lukyanov, S., Tissen, E., & Kislyak, N. (2007). The Market for Air Passenger Operations: Contestable Market or...? *Voprosy Ekonomiki*, 11, 120–138. <https://doi.org/10.32609/0042-8736-2007-11-120-138>

Markova, O. A., Meleshkina, A. I., & Morosanova, A. A. (2019). Tariff Setting within Airlines Competitive Strategies. *Scientific Research of Faculty of Economics. Electronic Journal*, 11(4), 19–34. <https://doi.org/10.38050/2078-3809-2019-11-4-19-34>

Matyukha, S. V. (2022). Evolution of the model of the regional air transportation market. Market consumers. *Transport business of Russia*, 4, 104–107.

Skryleva, E. V. (2018). Investigation of factors affecting the development of regional air transportation. *Innovation and investment*, 9, 236–240.

Приложение А

Таблица А.1

Метрики пробит-модели	
Метрика	Пробит
accuracy	0,993
f1-score	0,498
precision	0,497
recall	0,5

Источник: построено авторами.

Таблица А.2

Метрики при случайном ресемплировании			
	Логит	Случайный лес	Градиентный бустинг
Метрика			
accuracy	0,839	0,961	0,952
f1-score	0,462	0,496	0,494
precision	0,5	0,5	0,5
recall	0,495	0,502	0,503

Источник: построено авторами.

Таблица А.3

Метрики при использовании весов в функции потерь

	Логит	Случайный лес	Градиентный бустинг
Метрика			
accuracy	0,85	0,96	0,938
f1-score	0,467	0,494	0,488
precision	0,501	0,499	0,499
recall	0,524	0,495	0,49

Источник: построено авторами.

Таблица А.4

Метрики при SMOTE

	Логит	Случайный лес	Градиентный бустинг
Метрика			
accuracy	0,855	0,956	0,943
f1-score	0,468	0,49	0,487
precision	0,5	0,498	0,497
recall	0,509	0,487	0,48

Источник: построено авторами.

Таблица А.5

Метрики при ADASYN

	Логит	Случайный лес	Градиентный бустинг
Метрика			
accuracy	0,849	0,963	0,952
f1-score	0,466	0,497	0,493
precision	0,5	0,501	0,5
recall	0,506	0,503	0,497

Источник: построено авторами.

Приложение Б

Таблица Б.1

Коэффициенты логит-модели

	Веса	Ресемплирование	SMOTE	ADASYN
served_airports_pr_period	−0,18318	−0,08333	−0,21262	−0,14403
neigh_airports_pr_period	1,033457	2,626483	2,483097	2,281484
airport_type	0,177997	0,295872	0,022516	−0,10024
population	0,339319	0,397428	0,778479	0,825467
megapolis	−0,38987	−0,35678	−0,44555	−0,42313
tourist	0,404815	0,502083	0,288666	0,249608
distance_km	−0,40461	−0,3507	−0,38226	−0,45651

Примечание: переменные стандартизированы.

Источник: построено авторами.

Таблица Б.2

Важность переменных при построении случайного леса

	Веса	Ресемплирование	SMOTE	ADASYN
served_airports_pr_period	0,149107	0,1556	0,132742	0,144254
neigh_airports_pr_period	0,230763	0,244797	0,223677	0,203452
airport_type	0,039299	0,039757	0,029445	0,032212
population	0,262579	0,25002	0,275388	0,278981
megapolis	0,033917	0,02838	0,036999	0,034832
tourist	0,093142	0,088691	0,079987	0,069684
distance_km	0,191193	0,192754	0,221763	0,236586

Примечание: под важностью понимается способность независимой разграничивать значения объясняемой переменной.

Источник: построено авторами.

Таблица Б.3

Важность переменных при построении градиентного бустинга

	Беса	Ресемплирование	SMOTE	ADASYN
served_airports_pr_period	−0,15248	−0,03639	0,012794	0,00788
neigh_airports_pr_period	0,192012	0,070452	0,085841	0,089573
airport_type	−0,00068	−0,00479	0,008825	0,012001
population	−0,05011	−0,03695	0,091211	0,148322
megapolis	−−0,0281	0,013063	0,068939	0,073407
tourist	0,00578	−0,01078	0,007389	0,00362
distance_km	−0,02778	−0,01021	−0,00182	0,018682

Примечание: под важностью понимается то, насколько использование переменной снижает значение функционала ошибки.

Источник: построено авторами.