

Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Санкт-Петербургский государственный
экономический университет»

На правах рукописи

РЯБЧИКОВА ДАРЬЯ АНАТОЛЬЕВНА

**ПРОГНОЗИРОВАНИЕ ЧИСЛЕННОСТИ НАСЕЛЕНИЯ
В СОВРЕМЕННОМ МЕГАПОЛИСЕ
(НА ПРИМЕРЕ САНКТ-ПЕТЕРБУРГА)**

Специальность 5.2.3. – Региональная и отраслевая экономика
(Бухгалтерский учет, аудит и экономическая статистика)

Диссертация на соискание ученой степени
кандидата экономических наук

Научный руководитель:
доктор экономических наук,
профессор, член-корреспондент РАН,
заслуженный деятель науки РФ
Елисеева И.И.

Санкт-Петербург – 2026

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	4
ГЛАВА 1. МЕГАПОЛИС КАК ОБЪЕКТ ДЕМОГРАФИЧЕСКОГО ПРОГНОЗИРОВАНИЯ	14
1.1. Мегаполис как открытая демографическая система.....	14
1.2. Демографические компоненты мегаполиса и подходы к прогнозированию.....	18
1.3. От глобальных сценариев к городским прогнозам: роль методологии ООН.....	24
1.4. Опыт Барселоны в прогнозировании демографической динамики крупного города.....	29
ГЛАВА 2. ОСОБЕННОСТИ ДЕМОГРАФИЧЕСКОЙ ДИНАМИКИ МЕГАПОЛИСА: УСЛОВИЯ И ОГРАНИЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ (НА ПРИМЕРЕ САНКТ-ПЕТЕРБУРГА).....	42
2.1. Компонентный профиль демографической динамики Санкт-Петербурга: баланс естественного и миграционного движения	42
2.2. Миграция как ведущий и наименее предсказуемый компонент воспроизводства мегаполиса	44
2.2.1. Динамика внутренней миграции.....	44
2.2.2. Социально-демографический профиль внутренних и внешних мигрантов	47
2.2.3. Пространственно-экономические детерминанты миграции: расширенная гравитационная модель	52
2.2.4. Статистические ограничения и нерегистрируемая мобильность	56
2.3. Рождаемость в условиях мегаполиса: тенденции и последствия для демографического прогноза	58
2.4. Смертность в мегаполисе: устойчивость показателей и значение возрастной структуры	62
2.5. Совокупное воздействие демографических компонентов и обоснование вероятностного подхода к прогнозированию	65

2.5.1. Подготовка данных о рождаемости	69
2.5.2. Подготовка данных о смертности	69
2.5.3. Подготовка данных по миграции	76
2.5.4. Стартовая численность населения	83
ГЛАВА 3. КОГОРТНО-КОМПОНЕНТНЫЙ ПРОГНОЗ ЧИСЛЕННОСТИ НАСЕЛЕНИЯ МЕГАПОЛИСА: ВЕРОЯТНОСТНЫЙ ПОДХОД.....	86
3.1. Методика вероятностного когортно-компонентного прогноза	86
3.1.1. Концептуальные основы	86
3.1.2. Стохастическое моделирование компонент (PCA-AR(1))	91
3.1.3. Стохастическая когортно-компонентная симуляция (Монте-Карло- модель).....	99
3.1.4. Калибровка предиктивных интервалов по результатам ретроспективного моделирования (backtest)	107
3.1.5. Построение итоговых предиктивных интервалов и визуализация результатов	108
3.2. Результаты базового вероятностного прогноза численности и возрастно- половой структуры населения Санкт-Петербурга.....	110
3.3. Перспективные траектории компонент воспроизводства населения.....	116
3.4. Сценарный анализ в вероятностном прогнозе населения Санкт- Петербурга.....	118
ЗАКЛЮЧЕНИЕ	132
СПИСОК ЛИТЕРАТУРЫ.....	135
ПРИЛОЖЕНИЯ.....	148

ВВЕДЕНИЕ

Актуальность темы диссертационного исследования.

В современном мире происходит «стягивание» населения в города: доля городского населения Земли увеличилась с 33,7% в 1960 г. до 57,7% в 2024 г.¹ Особую роль приобретают города-мегаполисы. В российской статистической практике отсутствует нормативно закреплённое определение мегаполиса. В научных работах по экономике и урбанистике представлены различные определения данного понятия. Так, в экономических исследованиях мегаполис рассматривается как форма пространственного развития, в рамках которой проявляются агломерационные эффекты, при которых более крупные города и регионы оказываются более эффективными и производительными по сравнению с меньшими территориями². В урбанистике мегаполис определяется как кластер хорошо связанных городов, формирующий единое функциональное пространство³.

К мегаполису, в российской практике, относят города с населением свыше одного миллиона человек («города-миллионники»), являющиеся центрами городских агломераций. Ряд исследователей считает [32], что численность населения является важным, но не определяющим признаком: статус мегаполиса формируется сочетанием функционального назначения, роли в национальных и региональных трансформациях, административно-столичных функций и особенностей пространственной организации. В международной классификации ООН (UN DESA⁴) мегаполисом признаётся городская агломерация с населением более 10 млн человек, с высокой плотностью расселения и устойчивыми экономическими, социальными и инфраструктурными связями⁵. Этому критерию в России формально соответствует лишь г. Москва. Однако г. Санкт-Петербург, с

¹ World Bank. Urban population (% of total population). World Development Indicators. Source: United Nations Population Division, World Urbanization Prospects. URL: <https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS>

² orig.: “larger cities and regions are more efficient and more productive than smaller ones.” Cottineau C. et al. Defining urban agglomerations to detect agglomeration economies (2016).

³ Orig.: “A cluster of well-networked cities is called a megalopolis.” Dobbins T. What's the Difference Between a Megacity, a Metropolis, a Megalopolis and a Global City? ArchDaily, 2018.

⁴ United Nations Department of Economic and Social Affairs

⁵ United Nations, Department of Economic and Social Affairs, Population Division. The World's Cities in 2025: Data Booklet. 2025. P. 2. Orig.: “cities with more than 10 million inhabitants are classified as megacities”. URL: https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/undesa_pd_2025_data-booklet_world_cities_in_2025.pdf

населением свыше 5,5 млн человек, статусом «окна в Европу» и «культурной столицы России», являющийся центром Северо-Западного Федерального округа, по своим демографическим и функциональным характеристикам также обоснованно может быть отнесен к мегаполисам. Выбор Санкт-Петербурга в качестве объекта исследования обусловлен тем, что при сопоставимом уровне демографического старения его миграционная компенсация выражена заметно слабее, чем в Москве. Миграционный прирост в расчёте на 1000 жителей в базовом 2023 г. составил в Санкт-Петербурге 1,7‰, в Москве 3,1‰⁶. Так что, для Санкт-Петербурга особенно значим анализ не только общей величины миграционного прироста, но и того, насколько он способен компенсировать естественную убыль и старение населения.

Взятые в совокупности демографические характеристики позволяют рассматривать мегаполис как *открытую демографическую систему*, развитие которой определяется нестабильным и трудно прогнозируемым миграционным процессом. Если на уровне страны численность прибывших и выбывших внутренней миграции уравниваются, то мегаполис живёт по иным законам: здесь внутренняя миграция может варьировать, естественная убыль практически никогда не компенсируется рождаемостью, и именно миграция берёт на себя роль главного механизма динамики численности населения. Влияние миграции усиливает асинхронность воспроизводственных процессов, а также дифференциацию населения, его чувствительность к социальным, экономическим и эпидемиологическим шокам. Становится все более очевидно, что традиционные сценарные модели, применяемые Росстатом для прогнозирования численности и состава населения оказываются недостаточными.

Классические подходы к прогнозированию населения, основанные на экстраполяции трендов и фиксированных сценариях, игнорируют стохастическую

⁶ Рассчитано по данным ЕМИСС. Миграционный прирост в расчёте на 1000 жителей определялся как отношение миграционного прироста населения за 2023 г. к численности постоянного населения соответствующего города на 1 января 2024 г., умноженное на 1000. Источники: «Миграционный прирост населения, человек, за год». URL: <https://fedstat.ru/indicator/61749>; «Численность постоянного населения – женщин по возрасту на 1 января, человек». URL: <https://fedstat.ru/indicator/33459>; «Численность постоянного населения – мужчин по возрасту на 1 января, человек». URL: <https://fedstat.ru/indicator/31548>.

природу демографических процессов мегаполиса [72]. Они недостаточно отражают изменчивость миграционной динамики и поведенческих параметров рождаемости и смертности. В результате прогнозы для мегаполисов, полученные по стандартным схемам, часто расходятся с реальностью, подрывая эффективность социальной политики, планирования жилищного строительства, развития инфраструктуры. В этих условиях особую актуальность приобретает совершенствование методологии прогнозирования численности и состава населения, ориентированной на отражение неопределённости и вариативности демографических процессов и повышающее чувствительность прогноза к реальному развитию мегаполиса.

Степень разработанности научной проблемы. В развитие когортно-компонентного метода внесли вклад А.Г. Волков (1931 – 2009), Л.Е. Дарский (1930 – 2001), Е.М. Андреев, Т.Л. Харьковская, С.В. Захаров и другие. Значительный вклад в понимание мировых тенденций и их сочетания с региональными особенностями внесли межстрановые исследования М.А. Клушта [20, 23] и историко-социологический подход к изучению демографических процессов в работах А.Г. Вишневецкого (1931 – 2021). Отдельное место занимают также работы М. Б. Денисенко, Е. Л. Сороко, И. Е. Калабихиной, Л. И. Ниворожкиной, Г. Л. Сафаровой, Н. В. Мкртчяна, М. Г. Носовой, О. С. Чудиновских и С. Ю. Никитиной, в которых рассматриваются вопросы статистического анализа, моделирования и прогнозирования социально-экономических и демографических процессов [16, 36, 41, 45, 48, 56, 57, 60].

Зарубежные исследования последних десятилетий содержат широкий спектр вероятностных решений [78, 112]. Ключевым этапом стала модель Р. Ли и Л. Картера (R. D. Lee, L. R. Carter) для прогнозирования смертности, в которую впервые была введена стохастическая компонента [93]. Последующие работы В. Лутца, В. Сандерсона и С. Щербова (W. Lutz, W. Sanderson, S. Scherbov) сформировали концепцию вероятностных прогнозов рождаемости и смертности [97]. Современные работы Р. Дж. Хайндмана, Р. Ли и Л. Тикла (R. J. Hyndman, R. Lee, L. Tickle), основанные на методах анализа временных рядов и

функциональных данных, расширили возможности моделирования [87]. В развитие вероятностного демографического прогнозирования также внесли вклад Н. Кейлман (N. Keilman), Д. Фам (D. Pham), А. Хетланд (A. Hetland), А. Е. Рафтери (A. E. Raftery), Н. Ли (N. Li), Х. Шевчикова (H. Ševčíková), П. Ванелла (P. Vanella), П. Дешермейер (P. Deschermeier) и К. Ю (K. Yu). Новейшие разработки включают байесовские и пространственно-временные модели, применимые к условиям мегаполисов [79, 88, 94, 96, 100, 103, 107].

В современной российской официальной статистической практике прогнозирование по-прежнему опирается на детерминистские сценарии, где рассчитываются низкий, средний и высокий варианты прогноза. С. Ю. Никитина обосновывает необходимость адаптации традиционных методов прогноза численности населения к условиям недостатка демографических данных, особенностям российской статистической базы и миграционной динамики [47]. На развитие методов анализа и моделирования существенное влияние оказали работы И. И. Елисеевой и её научной школы, в которых систематизированы статистические и эконометрические подходы к исследованию социально-экономических и демографических процессов [13].

Цель диссертационного исследования – обоснование и реализация метода статистического прогнозирования численности населения мегаполиса на основе вероятностного когортно-компонентного подхода с учётом стохастического характера рождаемости, смертности и миграции, исходя из специфики мегаполиса как открытой демографической системы (на примере г. Санкт-Петербурга).

Реализация поставленной цели потребовала решения выполнения **следующих задач:**

- уточнение специфики мегаполиса как открытой демографической системы и определение методологических ограничений традиционных сценарных моделей прогнозирования, применительно к крупным городам;
- критическое обобщение отечественных и международных подходов к демографическому прогнозированию мегаполиса, включая методику Росстата и систему прогнозов ООН; оценку их применимости к условиям мегаполиса;

- проведение сравнительного анализа демографической динамики Санкт-Петербурга и Барселоны как примеров крупных городов с различными моделями воспроизводства населения и миграционного регулирования и, соответственно, особенностями прогнозирования;
- разработка модифицированной методики прогнозирования на основе стохастического когортно-компонентного подхода, учитывающей неопределённость миграционных потоков и вариативность демографических процессов;
- обоснование выбора эмпирической базы для калибровки вероятностной модели и проверка чувствительности результатов к альтернативным историческим окнам;
- построение вероятностного прогноза численности и возрастно-половой структуры населения Санкт-Петербурга до 2045 года и сопоставление полученных оценок с официальным прогнозом Росстата;
- получение прогнозных оценок рождаемости, смертности и чистой миграции (включая суммарный коэффициент рождаемости, число родившихся и умерших, ожидаемой продолжительности жизни при рождении) и проверку чувствительности оценок к детализации возрастной структуры.
- разработка сценарного анализа результатов вероятностного прогноза на основе четырёх сценариев демографического развития Санкт-Петербурга – от благоприятного до неблагоприятного, включая два промежуточных варианта, и сопоставление полученных оценок с официальным прогнозом Росстата.

Объектом исследования выступает мегаполис как открытая демографическая система. В качестве эмпирического примера выбран Санкт-Петербург – второй по численности населения город России, сформировавший агломерацию, крупный центр внутренней и внешней миграции, обладающий особой демографической структурой.

Предметом исследования являются методологические и методические аспекты вероятностного когортно-компонентного демографического прогнозирования численности и состава населения мегаполиса, учитывающие

специфику изменений рождаемости, смертности, внутрироссийской и внешней миграции, что обеспечивает возможность вариативности стратегического планирования мегаполиса.

Теоретическую основу диссертационного исследования составили труды российских и зарубежных ученых в области демографического прогнозирования, теории воспроизводства населения, когортно-компонентного метода и вероятностного моделирования демографической динамики. Использованы научные положения и методические разработки по вопросам статистического анализа демографических процессов, моделирования рождаемости, смертности и миграции, построения прогнозных оценок численности и возрастно-половой структуры населения.

Методологическую основу исследования составили когортно-компонентная модель прогнозирования, вероятностное моделирование с использованием Монте-Карло-симуляций, методы анализа временных рядов, метод главных компонент, построение сценарных прогнозов, ретроспективная проверка прогноза и калибровка предиктивных интервалов [105].

Информационная база исследования. Диссертационное исследование выполнено на основе открытых данных Федеральной службы государственной статистики (Росстат), Единой межведомственной информационно – статистической системы (ЕМИСС), а также данных Управления Федеральной службы государственной статистики по Санкт-Петербургу и Ленинградской области (Петростат).

Обоснованность результатов исследования обеспечена опорой на концепции и методы современной демографической науки, статистики и демографического прогнозирования, включая представление о мегаполисе как открытой демографической системе. Достоверность и внутреннюю согласованность полученных выводов обеспечили применение когортно-компонентного подхода, вероятностного моделирования, методов анализа временных рядов, процедур вторичной группировки исходных данных, а также калибровки предиктивных интервалов на основе ретроспективной проверки

модели. Существенное значение имело обращение к международному опыту демографического прогнозирования, прежде всего к методологическим принципам ООН и зарубежной практике городских прогнозов, что позволило соотнести предложенный подход с современными стандартами вероятностного анализа демографической динамики.

Достоверность результатов исследования обеспечивается использованием официальных статистических данных Росстата, Петростата и рекомендаций ООН, использованием достижений отечественной и мировой науки, что делает возможным корректное сравнение полученных оценок с официальными прогнозами. Результаты исследования подтверждены апробацией на международных и всероссийских конференциях и публикациями результатов в рецензируемых научных изданиях, включая журналы из перечня ВАК, имеющих право публикации научных работ по специальности 5.2.3.

Соответствие диссертации паспорту научной специальности. Представленная научно-квалификационная работа соответствует паспорту специальности 5.2.3 – «Региональная и отраслевая экономика» (Бухгалтерский учет, аудит и экономическая статистика):

П. 11.15. Статистическое наблюдение. Методы сбора и обработки статистической информации.

П. 11.17. Прикладные статистические исследования в экономике. Статистическая поддержка управленческих решений.

Научная новизна диссертационного исследования заключается в обосновании и модификации вероятностного подхода к прогнозированию численности и состава населения мегаполиса на основе стохастического когортно-компонентного моделирования, позволяющего представить прогноз как спектр возможных демографических траекторий в стратегическом планировании и управлении развитием мегаполиса.

К числу наиболее значимых и обладающих новизной научных результатов, полученных лично соискателем, относятся следующие:

1. Обоснование прогнозирования населения мегаполиса как особого направления, учитывающего открытость мегаполиса как демографической системы;
2. Модификация вероятностного подхода к демографическому прогнозированию применительно к мегаполису, основанного на стохастическом когортно-компонентном моделировании, включающего построение прогнозной системы статистических показателей численности и состава населения;
3. Разработка процедуры подготовки и согласования эмпирической базы для вероятностного прогноза населения мегаполиса, включающей вторичную группировку демографических рядов, реконструкцию недостающих возрастно-половых показателей, а также проверку чувствительности результатов к альтернативным историческим окнам калибровки;
4. Обоснование стохастической спецификации PCA-AR(1) и процедуры калибровки предиктивных интервалов, что позволяет статистически обоснованно оценивать не только медианную траекторию прогноза, но и вероятный диапазон возможных демографических отклонений;
5. Построение долгосрочного вероятностного прогноза численности населения и возрастно-половой структуры населения Санкт-Петербурга до 2045 г., включающий оценку медианной траектории, предиктивных интервалов, перспективных параметров рождаемости, смертности и миграции.
6. Разработка перспективных сценариев демографического развития Санкт-Петербурга и оценка влияния сценарных демографических параметров на численность и возрастно-половую структуру мегаполиса в целях статистической поддержки управленческих решений и стратегического планирования развития мегаполиса.

Теоретическая значимость исследования заключается в развитии вероятностного подхода к демографическому прогнозированию мегаполиса. Результаты исследования расширяют теоретические представления о методах прогнозирования населения мегаполиса, как открытой системы в условиях стохастичности демографических процессов.

Практическая значимость работы состоит в том, что разработанная методика прогнозирования численности и состава населения Санкт-Петербурга может быть использована Росстатом для включения в официальную методику демографического прогнозирования для субрегионов. Результаты исследования могут быть использованы в стратегическом планировании развития российских мегаполисов. Разработанный подход способен повысить обоснованность и точность прогнозов населения, которые лежат в основе решений о развитии транспортной, жилищной и социальной инфраструктуры.

Апробация результатов исследования. Результаты исследования прошли всестороннюю апробацию на следующих научных конференциях:

- XX Международная научно-практическая конференция, "Проблемы социально-экономической устойчивости региона" (ПГАУ, Пенза, 25-26 января 2023);
- Евразийская олимпиада по теории статистики (ОГУ, Оренбург, 13-19 мая 2023);
- Всероссийская научно-практическая конференция (с международным участием) «Измерение и анализ благосостояния» (СПбГЭУ, Санкт-Петербург, 25-27 января, 2024)
- Научная сессия профессорско-преподавательского состава, научных сотрудников и аспирантов по итогам НИР «Актуальные вопросы развития современной науки: теория и практика» (СПбГЭУ, Санкт-Петербург, апрель-май, 2024)
- IX Санкт-Петербургский Международный Форум Труда, круглый стол «Кадровый потенциал регионов: диагностика, и прогнозирование, управление» (КВЦ «Экспофорум», Санкт-Петербург, 2-4 апреля, 2025)
- Всероссийская научная конференция «Серебряный возраст. Петербург» (Федеральный научно-исследовательский социологический центр Российской академии наук Социологический институт РАН – филиал ФНИСЦ РАН, Санкт-Петербург, 22 мая 2026)

Публикации результатов исследования. Основные положения и результаты исследования отражены в 8 научных статьях, в том числе в 4 статьях, опубликованных в рецензируемых журналах, включенных в рекомендованный список ВАК Российской Федерации, общим объемом 5,09 п. л., в том числе авторским – 4,72 п. л.

Структура диссертационного исследования. Цели и задачи исследования определили его структуру. Диссертация включает введение, три главы, заключение, список использованной литературы, кроме того, имеется десять приложений. Объём основного текста составляет 134 страницы, включая 16 таблиц, 6 листингов, 19 рисунков и 59 формул; список использованной литературы насчитывает 118 источников. Логика работы определяется последовательностью перехода от теоретического обоснования мегаполиса как особого объекта демографического прогнозирования к анализу демографической динамики Санкт-Петербурга, разработке вероятностной когортно-компонентной модели и интерпретации полученных прогнозных результатов.

ГЛАВА 1. МЕГАПОЛИС КАК ОБЪЕКТ ДЕМОГРАФИЧЕСКОГО ПРОГНОЗИРОВАНИЯ

1.1. Мегполис как открытая демографическая система

Мегполис представляет собой не только крупный по численности населения город, но и функциональное ядро агломерации, обладающее высокой экономической и демографической плотностью, развитой системой социально-экономических связей, способностью интегрировать периферию и формировать устойчивую зону влияния. По данным на 2023 год, в крупнейших мегполисах России (к ним относятся Москва, Санкт-Петербург, Новосибирск, Екатеринбург и Нижний Новгород [27]) сосредоточено почти 16%⁷ всего населения страны. На их долю приходится более 33%⁸ консолидированного валового регионального продукта, а миграционный прирост как правило, превышает среднерегиональные показатели.

С научной точки зрения мегполис целесообразно рассматривать не только как крупный по численности населения город, но и как функциональное ядро агломерации, обладающее высокой экономической и демографической плотностью, развитой инфраструктурой, выраженной социальной и пространственной дифференциацией [49, 83]. Мегполис – это вектор в эволюции урбанистической системы страны в соответствии с глобальными трендами [27]. Для уточнения объекта исследования разграничим понятия агломерации и мегполиса. Агломерация представляет собой совокупность населённых пунктов, объединённых трудовыми, транспортными, жилищными и социально-экономическими связями, тогда как мегполис выступает её институциональным и демографическим ядром. В случае Санкт-Петербурга такое разграничение особенно важно, поскольку город одновременно является самостоятельным субъектом статистического наблюдения и ядром агломерационной системы, связанной с ближайшими территориями Ленинградской области. В настоящей

⁷ Рассчитано автором по данным Росстата «Численность постоянного населения Российской Федерации по муниципальным образованиям на 1 января 2024 года (с учётом итогов Всероссийской переписи населения 2020 г.)» (дата обращения: 16.07.2025).

⁸ Рассчитано автором по данным ЕМИСС «Валовой региональный продукт в основных ценах (ОКВЭД 2)», URL: <https://fedstat.ru/indicator/61497> (дата обращения: 16.07.2025).

работе Санкт-Петербург рассматривается как мегаполис в административных границах; агломерационные связи с Ленинградской областью учитываются как фактор миграционной открытости, но не включаются в расчеты прогноза.

Мегаполис отличается демографической гетерогенностью и асинхронностью процессов: в нём могут одновременно сосуществовать районы с естественным приростом и депопуляцией, сегменты высокой рождаемости и отложенного родительства, центры привлечения мигрантов и зоны их вытеснения. За 2022 год чистое сальдо миграции в Санкт-Петербурге составило +6,571 тыс. человек, в то время как естественная убыль составляла -12,066 тыс. человек⁹. Особую нагрузку на инфраструктуру мегаполиса формирует маятниковая миграция [4]. В случае Петербурга речь идёт о регулярных ежедневных передвижениях населения между Ленинградской областью и городской агломерацией. Наиболее наглядно размытость границы между городом и областью проявляется в ближайших приграничных территориях Ленинградской области, включая г. Мурино, г. Кудрово, г. Бугры и дер. Новое Девяткино¹⁰. Формально они относятся к Ленинградской области, однако значительная часть повседневных трудовых, транспортных и потребительских связей их жителей ориентирована на Санкт-Петербург. По оценкам, в будний день в мегаполис въезжает до 110 тыс. человек¹¹. Эти перемещения не фиксируются в традиционных регистрах, фактически увеличивают нагрузку на транспорт, здравоохранение, образование и коммунальную сферу. Во многих странах развитие мегаполисов отражается в аналитических отчётах, где город рассматривается как целостная система – с демографией, экономикой, транспортом и социальной инфраструктурой. Отчёты City at a Glance (например, Сиднея), демографические профили Торонто и Парижа подробно характеризуют не только численность населения, но и ее компоненты.

⁹ ЕМИСС «Естественный прирост за год», URL: <https://fedstat.ru/indicator/31018>

¹⁰ Концепция комплексного развития территорий Ленинградской области, прилегающих к границам Санкт-Петербурга. Ленинградская область, 2014.

¹¹ Общий объем ежедневной маятниковой трудовой миграции, полученный в результате анализа данных о перевозках пассажиров пригородным железнодорожным транспортом для отдельных направлений по Ленинградской области, составил около 105 тыс. человек на 2013 г. (11% экономически активного населения региона). Бугаев, М. А. (2014). Маятниковые миграции на рынке труда Санкт-Петербурга и Ленинградской области. Вестник Санкт-Петербургского университета. Экономика, (4), 086-116.

Демографические прогнозы встроены в стратегическое планирование. Этот международный опыт показывает, почему для Санкт-Петербурга необходима методология прогноза, учитывающая его сложность как открытой демографической системы.

Комплексную структуру города отражает Ежегодный сборник Петростата «Санкт-Петербург». В отличие от зарубежных City reports, он ограничивается фиксацией текущего состояния и не содержит прогнозных сценариев.

Ключевой особенностью мегаполиса является рассогласованность фактической численности населения с официальными статистическими данными. Значительная доля мигрантов не регистрируется по месту пребывания, что приводит к систематическому занижению демографических показателей. В частности, реальное число студентов, временных работников, и иных временно присутствующих групп населения может существенно отличаться от учтённых в статистике значений.

Например, по данным независимых оценок, реальная численность населения Санкт-Петербурга в 2021 году могла превышать официально зарегистрированную на 1,5 миллиона человек¹², особенно в туристический сезон и периоды крупных событий. Это влечёт за собой ошибки в расчётах обеспеченности социальной инфраструктурой.

Кроме ежегодника публикуется бюллетень Петростата «Предположительная численность населения Санкт-Петербурга и Ленинградской области». В последнем издании бюллетеня содержится средний вариант прогноза Росстата до 2045 г.¹³ Для исследования этот прогноз важен прежде всего как официальный ориентир. Однако в бюллетене по каждому году представлено одно прогнозное значение. Хотя

¹² «Если по статистике у нас около 5,4 миллиона постоянных жителей (это данные 2019 года), то по нашим оценкам, движению людей, всему остальному мы получаем 7 миллионов постоянного населения. И, поверьте, анализ, который делают коллеги из налоговой, больше склоняется к этой цифре», – заявил глава Комитета по труду и занятости Петербурга Дмитрий Чернейко. Такие данные получены в результате исследования с применением технологии big data. URL: <https://www.sobaka.ru/city/city/138308>

¹³ Предположительная численность населения Санкт-Петербурга и Ленинградской области до 2045 года : статистический бюллетень / Петростат. URL: https://78.rosstat.gov.ru/storage/mediabank/%D0%9F%D1%80%D0%BE%D0%B3%D0%BD%D0%BE%D0%B7_2024%D0%A1%D0%9F%D0%B1_%D0%9B%D0%9E_14000524%20_%D0%B8%D1%82%D0%BE%D0%B3_1.pdf (дата обращения: 31.05.2026).

Росстат разрабатывает три сценарных оценки численности населения субъектов РФ и России в целом (низкая – средняя – высокая), но публикуется только средняя оценка. Оценить диапазон возможных отклонений по однозначным данным нельзя. Для Санкт-Петербурга особенно важен иной подход. Численность города зависит не только от рождаемости и смертности, но и от миграции, которая заметно меняется год от года.

Центр мегаполиса отличается неоднородной возрастной структурой. С одной стороны, заметна доля молодых возрастов. Население 20-34 лет составляет в Санкт-Петербурге 17,8% против 17,1% в среднем по России¹⁴. С другой стороны, аккумулируется высокая доля пожилого населения. Такое сочетание важно для прогноза, оно формирует более поздние браки, откладывание рождения детей и более низкую ориентацию на многодетность.

В Санкт-Петербурге, по данным на 2023 год, коэффициент обеспеченности местами в дошкольных организациях составил 852 места на 1 000 детей дошкольного возраста¹⁵. На первый взгляд показатель высок, однако за усреднённым значением скрывается неравномерность: в новых районах фиксируются очереди, часть учреждений работает на пределе мощности. Большая инерционность систем расселения и ограничения, налагаемые самим процессом жилищного строительства, дополнительно сдерживают агломерационные процессы [27]. Поэтому планирование жилья и социальной инфраструктуры требует актуальных и методологически обоснованных демографических оценок.

Применительно к Санкт-Петербургу, одному из двух крупнейших мегаполисов России, все эти характеристики приобретают особую значимость. Город является ядром Санкт-Петербургской агломерации, демонстрирует миграционный прирост, но сталкивается с проблемами демографического старения и пространственного расслоения. Эти факторы делают задачу адаптивного

¹⁴ Рассчитано автором по данным ЕМИСС, Численность постоянного населения – мужчин по возрасту на 1 января и Численность постоянного населения – женщин по возрасту на 1 января, URL: <https://fedstat.ru/indicator/33459>
<https://fedstat.ru/indicator/31548>

¹⁵ Данные ЕМИСС «Обеспеченность детей дошкольного возраста местами в организациях, осуществляющих образовательную деятельность по образовательным программам дошкольного образования, присмотр и уход за детьми, приходится мест на 1000 детей», URL: <https://fedstat.ru/indicator/59530>

прогнозирования особенно важной для устойчивого развития города и агломерации.

1.2. Демографические компоненты мегаполиса и подходы к прогнозированию

Прогнозирование численности населения мегаполисов невозможно без опоры на всесторонний анализ ключевых компонентов демографической динамики – рождаемости, смертности и миграции. Эти компоненты формируют базу демографического воспроизводства в городской среде. В условиях мегаполиса эти компоненты проявляются в усложнённой форме: фертильность становится более вариативной и чувствительной к социальному и этническому контексту, структура смертности в значительной мере отражает неравенство условий жизни, а миграция отличается масштабностью, неоднородностью и фрагментарным отражением в статистике.

Для построения прогноза численности населения мегаполиса недостаточно опираться лишь на текущие данные – необходимо учитывать методологическое наследие отечественной демографии. Исторически в отечественной науке формировались разные подходы к прогнозированию рождаемости, смертности и миграции, которые сегодня могут быть адаптированы к условиям мегаполиса.

Интерес к прогнозированию населения в России восходит к XVIII веку: М. В. Ломоносов подчёркивал стратегическую значимость народонаселения, а И. Ф. Герман и Д. И. Менделеев предпринимали первые попытки количественного предвидения, заложив основы демографической прогностики [34, 35]. В советский период благодаря усилиям С. Г. Струмилина, Б. Ц. Урланиса, Э. А. Араб-Оглы, Л. Л. Рыбаковского и других прогнозирование приобрело системный характер, ориентированный, в основном на задачу воспроизводства трудовых ресурсов, планирования социального обеспечения и развития инфраструктуры [2, 52, 61, 62].

Фундаментальные изменения воспроизводственных процессов, охватившие российское общество во второй половине XX века, обусловили необходимость переоценки прежних методик. А. Г. Вишневский и С. В. Захаров в своих исследованиях подчёркивают, что процессы урбанизации, демографического

перехода и изменения брачно-семейного поведения размывают прежние зависимости между компонентами демографической динамики и требуют пространственно-дифференцированного анализа [7]. Особое внимание уделяется тому, что мегаполисы становятся территориальными «лапшевыми бумажками» демографических трансформаций, где рождаемость снижается быстрее, смертность приобретает новую структуру, а миграция выступает не только компенсатором потерь, но и фактором внутреннего расслоения населения.

Для настоящего исследования важное значение имеет работа С. В. Никитиной, поскольку она ориентирована на прогнозирование в условиях недостатка демографических данных [47]. Для мегаполиса эта проблема особенно заметна. Миграционные потоки меняются быстрее, чем рождаемость и смертность, поэтому средние региональные или общероссийские тренды не всегда передают специфику развития города.

В работах И. И. Елисеевой показано, что решение о рождении ребёнка связано не только с установками семьи, но и с условиями её жизни, включая жильё, социальную поддержку, правовые и организационные ограничения [12, 14]. В крупном городе эти факторы сильнее влияют на сроки рождения детей и снижают вероятность перехода семьи к двум и более детям. Особенно важно, что в условиях мегаполиса происходит «разрыв» между репродуктивными установками и их реализацией, а значит, прогнозирование рождаемости должно учитывать не только демографические коэффициенты, но и институциональные барьеры.

С. В. Захаров в фундаментальном исследовании «История рождаемости в России: от поколения к поколению» рассматривает историческую динамику рождаемости через сопоставление показателей условных и реальных поколений [15]. Такой подход позволяет отделить краткосрочные колебания рождаемости в отдельные календарные годы от более устойчивых изменений репродуктивного поведения поколений, что важно для построения реалистичных демографических прогнозов.

В исследованиях Л. Л. Рыбаковского и О. Л. Рыбаковского подчёркивается, что миграционный прирост может как компенсировать, так и усугублять

последствия депопуляции на региональном уровне [51]. Для российских крупных городов эта проблема дополнительно связана с совмещением внутренней, международной и временной трудовой миграции, а также с перераспределением населения между городами и пригородами [17, 38, 39]. В зарубежной демографической литературе для описания возрастных профилей миграции применяются модельные миграционные расписания, позволяющие сглаживать нестабильные возрастные ряды, учитывать студенческие пики мобильности и использовать миграционные профили в прогнозных расчётах [102; 111]. При этом сама структура миграции, по наблюдениям В. В. Локосова, становится всё более гетерогенной: усиливается доля временных и сезонных мигрантов, нерегистрируемых перемещений и маятниковой мобильности [33]. Эти формы миграции плохо поддаются статистическому наблюдению, что создаёт проблемы в построении надёжных моделей и требует привлечения альтернативных источников данных – сотовой аналитики, данных валидаторов, регистраций в электронных сервисах и др. В работах Н. М. Римашевской показано, что прогноз рождаемости не может опираться только на средние показатели, поскольку семьи различаются по условиям жизни, установкам и возможностям реализации репродуктивных намерений [40, 42, 50]. Эта мысль близка к выводам М. А. Клупта о «хрупкости репродуктивной функции семьи» в развитых странах. Семейно-демографическая политика должна учитывать не усреднённую модель семьи, а различие ценностных установок и повседневных практик [24].

Для Москвы и Санкт-Петербурга, как мегаполисов, эти вопросы подробно рассматриваются в работах Е. И. Кубасовой [29-31]. По её данным, с начала 1990-х годов в обеих столицах меняется распределение рождений по возрасту матери, растёт значение отложенных рождений и сокращается доля молодых матерей. На репродуктивное поведение всё заметнее влияют образование, занятость и жилищные условия.

В исследованиях Т. Л. Харьковской и Е. М. Андреева показано, что смертность в российских городах имеет устойчивые отличия от зарубежных городских популяций [1, 64]. Особенно различны причины смерти в трудоспособных

возрастах, включая внешние причины, болезни органов пищеварения и алкогольно-ассоциированные заболевания [1, 64]. Для прогноза это существенно, поскольку такие потери по-разному влияют на возрастную структуру населения и трудовой потенциал города.

Исследования О. С. Чудиновских [66, 68], Н. Е. Чистяковой [65], В. М. Федорищенко [63] и обращают внимание на проблему неучтённой миграционной мобильности. Маятниковые перемещения и краткосрочное пребывание населения искажают оценку текущей численности и, как следствие, расчёты потребности в транспорте, здравоохранении и социальной инфраструктуре. В работах О. С. Чудиновских подчёркивается, что интерпретация миграционной статистики должна учитывать не только масштабы перемещений, но и методологию их регистрации, поскольку изменение критериев учета влияет на объёмы и структуру фиксируемых миграционных потоков [66; 68]. Поэтому официальную статистику целесообразно дополнять другими источниками данных. При этом, как отмечают М. А. Клупт и О. Н. Никифоров, такая работа требует прозрачной методологии и предварительной качественной оценки данных из разных источников [22, 67].

В качестве основной государственной методики в России используется модель Росстата, основанная на компонентном подходе и системе сценарных гипотез [9]. Прогноз составляется в трёх вариантах – низком, среднем и высоком – и реализуется через процедуру «передвижки по возрастам», учитывая динамику рождаемости, смертности и миграции.



Рис. 1. Блок-схема взаимосвязанных прогнозных расчетов на национальном и региональном уровнях.

Источник: методология Росстата «Демографический прогноз численности населения до 2046 года»

Подробный алгоритм реализации прогнозных расчётов Росстата представлен в виде блок-схемы (рис. 1), демонстрирующей этапы сценарной проработки, независимых расчётов по субъектам, а также процедуру балансировки возрастных компонент на национальном уровне.

Прогноз Росстата строится в трёх сценариях – низком, среднем и высоком, однако для субъектов РФ, в том числе для Санкт-Петербурга, в открытом доступе публикуется только средний вариант прогноза. Он не сопровождается анализом неопределённостей и диапазоном вероятных траекторий, что ограничивает его ценность для стратегического планирования. В международной статистической практике демографические проекции трактуются не как единственное предсказание будущего, а как условные сценарии, основанные на предположениях о рождаемости, смертности и миграции [99].

Согласно среднему варианту прогноза Росстата, численность населения Санкт-Петербурга плавно снижается до 5,5 млн человек, несмотря на сохранение миграционного притока и активное жилищное строительство (рис. 2).

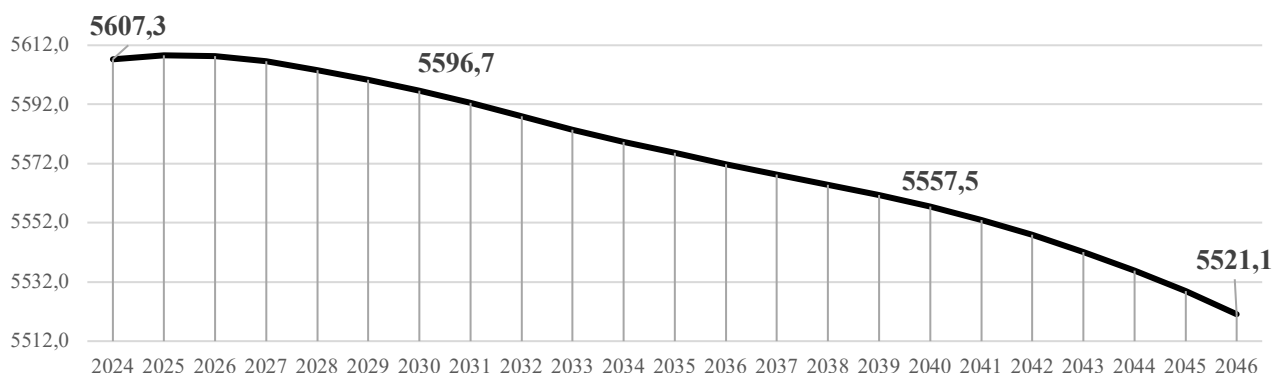


Рисунок 2 – Динамика прогнозной численности населения Санкт-Петербурга по прогнозу Росстата, 2024 – 2046 гг. (на начало года)

Источник: Статистический бюллетень «Предположительная численность населения Российской Федерации до 2045 года по среднему варианту демографического прогноза», СПб: Петростат, 2024.

Такая динамика показывает, что миграционный прирост не перекрывает влияние низкой рождаемости и старения населения и в официальном прогнозе эти тенденции экстраполируются. Для Санкт-Петербурга прогноз важен не только с точки зрения общей численности, но и с точки зрения будущей возрастной структуры. Более реалистичным был бы прогноз представленный не одной расчётной линией, с учетом возможного диапазона отклонений.

Отмеченные несоответствия требуют переосмысления методологии прогноза применительно к мегаполису как к открытой системе, чувствительной к множеству внешних и внутренних воздействий. Такое переосмысление особенно важно в условиях кризисов, поскольку, как отмечает М. А. Клупт, «стадиальные теории демографического развития, предназначенные для осмысления длительных и плавных процессов, теряют в этот период объяснительную и прогностическую силу» [19]. Необходим не только пересмотр параметров расчёта, но и переход к подходам, моделирующим вариативность будущего – с учётом неопределённости миграционных потоков, поведенческой изменчивости населения и уязвимости городской системы к шокам.

В данном исследовании предпринята попытка соединить российскую демографическую традицию с вероятностными методами, широко применяемыми

в мировой практике. Такой прогноз не только фиксирует тенденции, но и служит инструментом активного управления развитием мегаполиса.

1.3. От глобальных сценариев к городским прогнозам: роль методологии ООН

Разработка прогноза численности населения мегаполиса требует обращения не только к национальным, но и к международным источникам методологии. Среди них особое значение имеют руководства Организации Объединённых Наций, на протяжении десятилетий задающие стандарты демографического анализа и прогнозирования. Несмотря на ориентированность рассматриваемых моделей преимущественно на национальный уровень, они представляют собой методологическую основу, элементы которой могут быть адаптированы к условиям субнациональных образований – включая крупнейшие города и агломерации. Настоящий раздел посвящён характеристике системы прогнозирования ООН и оценке её применимости в решении задач демографического моделирования мегаполиса. Для мегаполисов, обладающих высокой миграционной текучестью и сложной внутренней структурой, при прогнозировании требуется учитывать динамику открытых систем. Международные подходы, задавая общие стандарты, не дают готовых решений, а формируют основу для адаптации к локальному уровню, включая уровень города. Значение международной сопоставимости городских данных подчёркивается в материалах ООН, ОЭСР и Евростата, где город рассматривается не только как административная единица, но и как функциональная территория, для которой важны единые критерии статистического наблюдения и сопоставимости показателей [82, 98, 113, 116].

Для мегаполисов, таких как Санкт-Петербург, обращение к международным методикам имеет не только академический, но и практический смысл. Эти подходы задают рамку, в которой можно сопоставлять городские прогнозы с глобальными трендами, а также адаптировать лучшие практики к условиям открытой, динамичной системы мегаполиса.

Система демографических прогнозов, разработанная Организацией Объединённых Наций и её статистическими подразделениями, представляет собой не только фундамент глобальной демографической аналитики, но и методологическую опору, на которую ориентируются национальные правительства, научные учреждения и муниципальные органы управления. Одним из ключевых продуктов, выпускаемых Отделом народонаселения Департамента по экономическим и социальным вопросам ООН (UN DESA), является доклад *World Population Prospects (WPP)* [114]. В его основе реализуются два типа прогнозов: детерминистские сценарии (*Standard projections*) и вероятностные траектории (*Probabilistic projections*), формируемые методом Монте-Карло с доверительными интервалами. WPP тем самым сочетает оба подхода. При этом вероятностные прогнозы разрабатывались и в других исследованиях, например в работе В. Щербова, В. Лутца и В. Сандерсона [97].

WPP публикуется каждые два года и включает оценки и прогнозы численности населения для всех стран и регионов мира, охватывая период с 1950 года до конца текущего столетия. Прогнозы предоставляются по полу и возрасту, с разбивкой по пяти- и однолетним возрастным группам, и включают как исторические данные, так и сценарии будущей рождаемости, смертности и миграции. Основу для построения этих прогнозов составляют национальные данные, переписи, выборочные обследования и регистрационные системы, прошедшие критическую оценку и сопоставление. Таким образом, система WPP служит основой для оценки демографических трендов в глобальном, региональном и страновом разрезах и формирует концептуальные рамки для разработки долгосрочных стратегий, выработки социальной и миграционной политики, а также для формирования локальных и региональных подходов, включая городские и мегаполисные прогнозы. Они воплощают в себе передовые достижения количественной демографии, выступая своего рода методическим эталоном в мировой практике. Однако для городов с высокой пространственной мобильностью населения эта «эталонность» становится условной и требует пересмотра ключевых допущений модели.

WPP изначально строится для стран, где миграция в основном международная и учтённая, а внутренняя мобильность менее значима. В мегаполисах же, особенно в российских, именно внутренняя миграция, маятниковые перемещения и временное население оказывают решающее влияние на демографическую динамику. В системе WPP миграционные потоки рассматриваются преимущественно как межстрановые, что соответствует национальному уровню прогнозирования, но требует адаптации при переходе к мегаполисам.

Центральное место в системе демографических прогнозов ООН занимает когортно-компонентный метод, применяемый для расчётов численности населения по полу и возрасту. В адаптации к субнациональному уровню его базовое балансовое уравнение отражает изменение численности населения под воздействием рождаемости, смертности, международной и внутренней миграции в течение прогнозного периода:

$$P_{(t+n)} = P_{(t)} + B_{(t \rightarrow t+n)} - D_{(t \rightarrow t+n)} + M_{(t \rightarrow t+n)}^{int} + M_{(t \rightarrow t+n)}^{dom}, \quad (1.1)$$

где t – начало прогнозного интервала, n – его длительность (в годах), $P_{(t)}$ – численность постоянного населения на момент t ; $P_{(t+n)}$ – численность постоянного населения на момент $t + n$; $B_{(t \rightarrow t+n)}$ – число родившихся за интервал $[t, t + n]$, $D_{(t \rightarrow t+n)}$ – число умерших за интервал $[t, t + n]$, $M_{(t \rightarrow t+n)}^{int}$ – чистая международная миграция за интервал $[t, t + n]$, $M_{(t \rightarrow t+n)}^{dom}$ – чистая внутренняя миграция за интервал $[t, t + n]$.

В такой форме уравнение (1.1) сохраняет логику когортно-компонентного метода и позволяет адаптировать его к субнациональному уровню, где внутренняя миграция становится самостоятельным фактором изменения численности населения.

В методологическом докладе World Population Prospects 2024 подчёркивается, что прогнозные ряды на 2024-2100 гг. строятся на основе вероятностных моделей суммарной рождаемости, ожидаемой продолжительности жизни и чистой международной миграции, которые затем интегрируются в

когортно-компонентную схему [115]. Опыт разработки демографических прогнозов для постсоветских стран показывает, что миграция остаётся наиболее сложной для задания компонентой прогноза, тогда как рождаемость и смертность обладают большей инерционностью [8]. Сходный вывод представлен и в работах по байесовскому прогнозированию миграции: миграционные потоки характеризуются высокой неопределённостью и слабой предсказуемостью в долгосрочном горизонте [70, 75, 76].

Дополнительную сложность создает разрыв между зарегистрированным и фактически проживающим населением, особенно в условиях высокой миграционной текучести. В статистических системах, использованных в WPP, в основу положены данные регистрации, тогда как реальные демографические процессы в крупных городах зачастую зависят от нерегистрируемых или временных форм присутствия. Это ведёт к смещению базовых показателей и снижает точность прогностических сценариев на локальном уровне.

Обновления WPP повышают детализацию национальных прогнозов, но при переходе к уровню мегаполиса требуют адаптации. В редакции 2022 года ООН отказалась от прежней схемы расчёта по пятилетним возрастным группам и пятилетним периодам и перешла к однолетним возрастам и годовым интервалам. Одновременно была уточнена внутренняя структура расчётов: возрастной диапазон расширен до 130 лет, опубликованные оценки и прогнозы приведены к верхней группе 100 лет и старше, а численность населения пересчитана на 1 января. Показатели рождаемости, смертности и миграции стали задаваться по календарным годам. Такая детализация делает прогноз более чувствительным к изменениям по возрасту и времени, однако для мегаполиса её всё равно необходимо дополнительно адаптировать с учётом миграционных особенностей. Кроме того, показатели рождаемости, смертности и миграции стали учитываться не за условные пятилетия, а по календарным годам, что обеспечило более высокую временную точность. Эти изменения отражены в сравнительной таблице (табл. 1), которая демонстрирует переход от традиционной схемы WPP к более детализированной структуре прогнозирования.

Таблица 1 – Изменения в реализации ККМПН между редакциями World Population Prospects за 2019 и 2022 годы

Рамки	2019 пересмотр	2022 Редакция
	5-летние возрастные группы и пятилетние периоды (5x5)	Однолетние возрастные группы и однолетние периоды (1x1)
Возрастной диапазон	От 0 до 110+	От 0 до 130+
Учетные даты численности населения	1 июля	1 января
Учетные периоды показателей рождаемости	1 июля по 30 июня	1 января по 31 декабря
Предположение о миграции	Конец периода или за период	Конец периода
Воздействие на население	Не вычисляется	Вычисляется

Источник: Methodology of the United Nations population estimates and projections «World Population Prospects 2022», URL: <https://population.un.org/wpp/>

На практике внедрение этих подходов осуществляется через итерационный процесс межпереписной верификации, представленный на рисунке 3. Механизм включает цикл уточнений и сопоставлений между расчётными данными и переписными оценками по полу и возрасту, с возможностью корректировки компонентов (рождаемости, смертности, миграции) при расхождении с бенчмарк-оценками. Такая процедура обеспечивает согласование данных в условиях высокой демографической изменчивости и нестабильных миграционных потоков, особенно актуальных в крупных городах и мегаполисах.



Рисунок 3 – Процесс, используемый для обеспечения межпереписной согласованности между демографическими компонентами и общей численностью населения.

Источник: The 2022 Revision of World Population Prospects <https://population.un.org/wpp/>

WPP рассчитан на национальный уровень, поэтому внутренняя территориальная мобильность не является самостоятельной компонентой прогноза. При переходе к субнациональному уровню, особенно в случае мегаполисов, встают вопросы корректности экстраполяции глобальных моделей на специфические локальные системы. Здесь стандартные инструменты, выработанные для стран, теряют точность, а ключевые допущения модели – о стабильности трендов и репрезентативности данных – оказываются не вполне реализуемыми. При переходе от странового прогноза к уровню мегаполиса ключевым ограничением становится внутренняя миграция. В WPP она напрямую не моделируется, поскольку система нацелена на национальный уровень. Для Санкт-Петербурга это принципиально, так как численность города зависит не только от рождаемости и смертности, но и от внутреннего притока мигрантов, маятниковых перемещений и временного присутствия населения в городе. Эти процессы могут быть учтены в пределах официальной статистики миграционного движения. В перспективе расширится использование данных мобильных операторов и геоинформационных данных для уточнения городских прогнозов.

Методология ООН задаёт общую компонентную логику прогноза, но для мегаполиса требует адаптации. Страновую модель нельзя механически перенести на город, где сильнее проявляются миграция, локальная возрастная структура и различия внутри территории.

1.4. Опыт Барселоны в прогнозировании демографической динамики крупного города

Обращение к опыту Барселоны связано не только с её статусом города-побратима Санкт-Петербурга. Барселона сопоставима с Санкт-Петербургом по роли крупного агломерационного ядра и располагает развитой системой городской демографической статистики. Поэтому её опыт используется как эмпирический

ориентир для анализа того, как крупный город учитывает миграцию, старение населения и различия внутри города.

Ещё в 1981 году было заключено соглашение о сотрудничестве между университетами Ленинграда и Барселоны, а в 1985 году подписан официальный Протокол о дружбе и взаимодействии. С тех пор практики стратегического планирования, городского управления, а также вопросы пространственного и демографического развития регулярно становились предметом двустороннего обмена.

В 1997 году делегация специалистов из Санкт-Петербурга изучала подходы Барселоны к долгосрочному городскому планированию, а в 2014 году петербургская сторона приняла участие во Всемирном конгрессе «Умный город», организованном ООН-Хабитат. В ходе мероприятий обсуждались вопросы миграционного управления, цифровизации городской среды, экологической устойчивости. Как подчеркнул А. В. Прохоренко, экс-глава Комитета по внешним связям Санкт-Петербурга, оба города «близки по духу, амбициям и темпам развития», а их партнёрство открывает возможности для внедрения проверенных решений [43, с. 8].

Барселона представляет собой пример мегаполиса, сумевшего в течение трёх десятилетий осуществить трансформацию городской среды, повысить конкурентоспособность экономики и добиться устойчивого роста численности и качества населения [28]. По ряду ключевых характеристик – от численности населения до территориального устройства – Барселона и Санкт-Петербург находятся в сопоставимом диапазоне.

Разработка первого стратегического плана развития (1997) ориентировалась именно на модель Барселоны, особенно в части интеграции городского пространства и управления миграцией. Однако при внешнем сходстве траекторий, характер демографических изменений в двух городах различается. Такое различие важно рассматривать не только как результат статистических расхождений, но и как проявление неодинаковых институциональных и ценностных контекстов демографического развития. В более широком теоретическом плане М. А. Клупт

отмечает, что в современном мире наблюдается «не глобальная вестернизация ценностей, а непрекращающийся ценностный конфликт» в вопросах семьи, рождаемости и миграции [21].

В количественном отношении оба города имеют схожий демографический масштаб: по данным на 1 января 2024 г. в Санкт-Петербурге проживало 5,598 млн. человек¹⁶, в агломерации Большой Барселоны – 5,805 млн. человек¹⁷. Демографическая траектория Барселоны наглядно демонстрирует, что устойчивый миграционный прирост в сочетании со сбалансированной возрастной структурой способен нивелировать последствия естественной убыли. Для Петербурга характерен демографический застой с умеренным приростом за счёт миграции, компенсирующей устойчивую естественную убыль. Барселона демонстрирует более устойчивый рост населения, поддерживаемый положительным миграционным сальдо, более стабильной рождаемостью и активной городской политикой. Различия заметны и в половой структуре: в 2023 г. в Барселоне мужчины составили 48,8%, женщины 51,2%, тогда как в Санкт-Петербурге, соответственно, 45,1% и 54,9%¹⁸.

Возрастная структура также различается. В Санкт-Петербурге выражен пик в группе 30-44 лет, связанный с демографическими волнами позднесоветского периода (рис. 4). Эта когорта формирует значительную часть трудоспособного населения и влияет на будущую нагрузку на рынок труда. В Барселоне распределение по этим возрастам более плавное.

¹⁶ По данным ЕМИСС «Численность постоянного населения на 1 января», URL: <https://fedstat.ru/indicator/31557>

¹⁷ По данным Испанского статистического института (National Statistics Institute) https://www.ine.es/jaxiT3/Datos.htm?t=2861#_tabs-tabla

¹⁸ По данным ЕМИСС «Численность постоянного населения – мужчин по возрасту на 1 января» и «Численность постоянного населения – женщин по возрасту на 1 января»

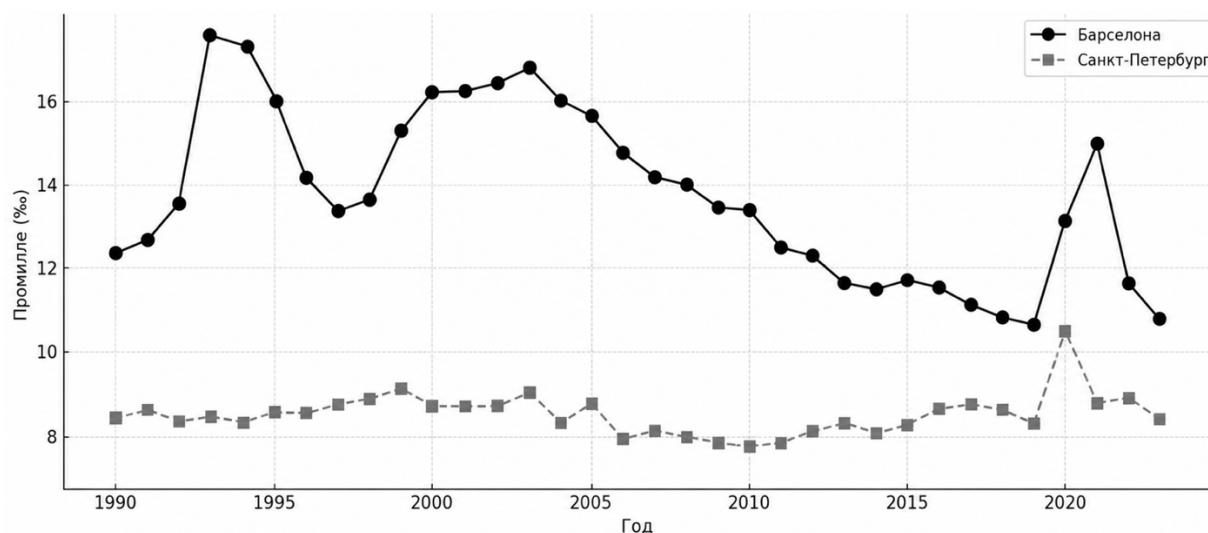


Рисунок 4 – Возрастная структура населения
в Санкт-Петербурге и Барселоне в 2023 году

Источник: составлено автором по данным ЕМИСС: «Структура численности постоянного населения на начало года (на 1 января) по полу и возрастным группам». URL: <https://fedstat.ru/indicator/43219> ; Instituto Nacional de Estadística (INE): Estadística del Padrón Continuo. Población por sexo, municipios y edad (año a año), municipio Barcelona, 2023.

Доля детей и подростков 0-14 лет в Барселоне выше, чем в Санкт-Петербурге, что может быть связано как с более высокой рождаемостью, так и с притоком молодых семей. В Санкт-Петербурге младшие возраста представлены слабее, что указывает на снижение числа рождений и ограниченный миграционный приток детей. Поэтому при прогнозировании для Санкт-Петербурга важно учитывать поколенческие волны, уже заложенные в возрастной структуре, а не только экстраполировать наблюдаемые коэффициенты рождаемости, смертности и миграции.

Динамика возрастной структуры оборачивается более высокой демографической нагрузкой на работающих петербуржцев. Для Барселоны характерна более «плоская» кривая старших возрастов, что связано не только с высокой продолжительностью жизни, но и большей ротацией населения. В Петербурге сохраняется высокая доля пожилых, не компенсированная миграцией в трудоспособном возрасте и усугубляющая уровень смертности.

Обе мегаполиса в 2023 году оказались в зоне естественной убыли населения, однако масштаб потерь различен: в Санкт-Петербурге: -11 832 человек, тогда как в

Барселоне: -9 690. Эти цифры подчёркивают общую тенденцию старения и замедление темпов воспроизводства, но также указывают на более острый характер демографического спада в российском мегаполисе.

Долгосрочная динамика общего коэффициента рождаемости демонстрирует различие траекторий демографического поведения в двух мегаполисах. В Санкт-Петербурге наблюдается выраженное падение рождаемости в начале 1990-х годов – до уровня 6,2 ‰, что отразило глубокий демографический кризис постсоветского периода. Лишь с 2000 года начался постепенный рост рождаемости, однако значения оставались ниже тех, что зафиксированы в Барселоне. В этом городе, динамика общего коэффициента рождаемости была более сглаженной, чем в Санкт-Петербурге: за 1990-2023 гг. среднее значение показателя составило 8,58‰. Вместе с тем показатель не был устойчивым на всём интервале и демонстрирует тенденцию к снижению: после максимума 9,69‰ в 2008 г. он снизился до 6,81‰ к 2023 г.¹⁹(рис. 5).

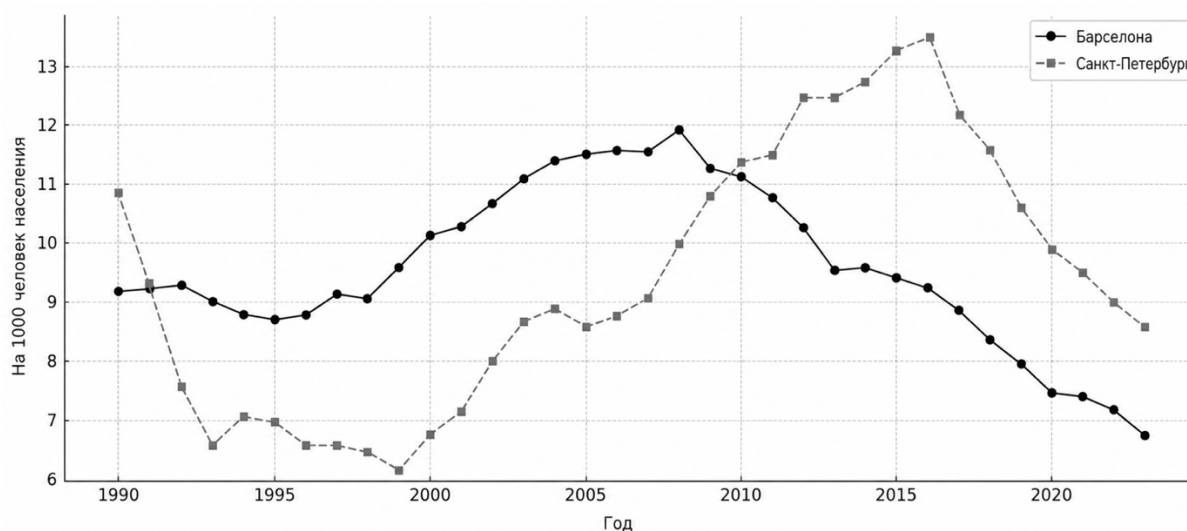


Рисунок 5 – Динамика общего коэффициента рождаемости
Санкт-Петербурга и Барселоны в 1990 – 2023 гг.

Источник: составлено автором по данным ЕМИСС: показатель «Число родившихся на 1000 населения за год». URL: <https://fedstat.ru/indicator/31269> ; Idescat: Taxa bruta de natalitat. Barcelona. URL: <https://www.idescat.cat/pub/?f=ssv&geo=mun%3A080193&id=inddt&n=917>

¹⁹ Idescat: Moviment natural. Fecunditat. Barcelona. Taxa bruta de natalitat. URL: <https://www.idescat.cat/pub/?f=ssv&geo=mun%3A080193&id=inddt&n=917>

Динамика общего коэффициента смертности в Санкт-Петербурге и Барселоне наглядно демонстрирует демографические издержки переходного периода 1990-х годов (рис. 6). В то время, как испанский мегаполис уже к началу 2000-х стабилизировал показатели рождаемости на уровне 10‰, Петербург ещё долго сохранял рецидивы постсоветской демографической турбулентности. Здесь смертность оставалась заведомо превышающей европейский уровень, фиксируя в отдельные годы значения выше 16‰. В последующие годы разрыв в уровне смертности сохранялся: в 2023 году – 10,7‰ в Петербурге против 8,4‰ в Барселоне. Причины лежат не только в различиях возрастной структуры, но и в более глубинных характеристиках: высокой распространённости хронических заболеваний, неоднородности социальной среды, запаздывающей реформе системы здравоохранения и меньшей защищённости уязвимых возрастных и социальных групп, климатических различиях. На этом фоне Барселона выступает примером мегаполиса, сумевшего выстроить сбалансированную систему демографической защиты – с акцентом на превентивную медицину, долгожительство и городскую среду, адаптированную к стареющему населению. Для прогноза населения это означает, что корректное моделирование смертности должно учитывать не только возрастную структуру, но и социально-медицинские различия, которые в крупных городах формируют специфические траектории смертности. Игнорирование этих факторов приводит к систематическим ошибкам прогноза. В прогнозной модели для Петербурга необходимо заложить чувствительность к различиям внутри города, что требует не только агрегированных данных, но и анализа внутригородских различий смертности, сигнализирующих о социальном неравенстве населения.

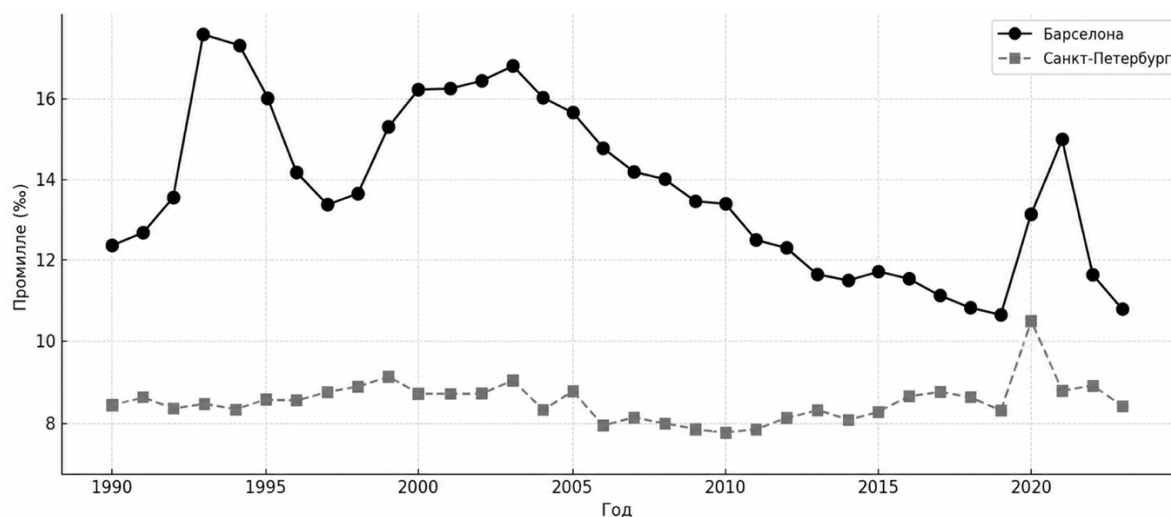


Рисунок 6 – Динамика общего коэффициента смертности в Санкт-Петербурге и Барселоне (1990 – 2023 гг.)

Источник: составлено автором по данным ЕМИСС: показатель «Число зарегистрированных умерших в расчете на 1000 населения (оперативные данные)». URL: <https://fedstat.ru/indicator/33537> ; Barcelona Datastore: Crude mortality rate (%). URL: <https://portaldades.ajuntament.barcelona.cat/en/statistics/39vcq6eho2> .

Разрыв в показателях миграционного прироста между Барселоной и Санкт-Петербургом трудно переоценить. Столица Каталонии привлекла за 2023 год почти 100 тысяч новых жителей за счёт миграции, в то время как за этот же год прирост в Петербурге составил лишь около 10 тысяч человек. Этот контраст иллюстрирует не только различие в масштабах международной и внутренней мобильности, но и особенности миграционной политики. В случае Барселоны миграция играет ключевую компенсаторную роль на фоне умеренного естественного убыли, тогда как в Петербурге её эффект существенно ограничен, в какой-то степени это результат латентной и маятниковой мобильности, не всегда попадающей в официальные учёты. Миграция в мегаполисе выступает не просто источником прироста, но и механизмом обновления возрастной структуры. Практика Испании показывает, что для прогноза важен не только общий объём миграции, но и её состав. Необходимо учитывать возраст мигрантов, направление потоков. Учитывая выявленные особенности демографической динамики Барселоны – положительное миграционное сальдо, устойчивую рождаемость и сбалансированную возрастную

структуру, представляет интерес то, как такие параметры учитываются в её прогнозных моделях.

В Испании прогнозирование населения опирается на компонентный метод и дополняется вероятностными и мультирегиональными элементами. Это позволяет учитывать различия между территориями и движение населения между ними. В 2022 г. Национальный институт статистики Испании опубликовал методологический бюллетень «Демографические прогнозы Испании 2022-2072 гг.», где представлены основные принципы построения долгосрочного прогноза.

В бюллетене чётко разделяют два подхода к построению прогноза. Метод *forecast* фокусируется на наиболее вероятном развитии событий, опираясь на текущие тенденции и статистически обоснованные ожидания. Этот метод применим, прежде всего, к краткосрочным горизонтам. Напротив, метод *project* предлагает построение сценариев при условии реализации заданных гипотез, даже маловероятных. Это позволяет исследовать последствия внешних шоков и чувствительность системы к различным параметрам.

Прогнозы Национального института статистики Испании, отражённые в бюллетене «Demographic Projections 2022-2072» [106], построены по методу *project*. Такой подход не стремится «угадать» точное будущее, но позволяет моделировать возможные сценарии на основе компонент рождаемости, смертности и миграции.

Результаты предоставляются на двухгодичной основе с горизонтом прогнозирования в 50 лет для всей страны и 15 лет для субъектов государства в соответствии с основными демографическими характеристиками (пол, возраст и поколение). По последним данным, итогами прогнозирования является численность населения, проживающего в Испании по состоянию на 1 января каждого года в период 2022-2072 гг., а также численность населения, проживающего в каждом из автономных сообществ и субъектов по состоянию на 1 января каждого года с 2022 г. по 2037 гг. В прогнозах предоставлены демографические события (рождения, смерти и миграционные движения), которые привели к эволюции объема и структуры населения в каждой из географических областей. Показатели как численности населения, так и его движения разбиваются

по основным демографическим характеристикам, таким как пол, возраст и год рождения (поколение). Для Санкт-Петербурга было бы ценно использование такого подхода, учитывая имеющуюся неравномерность замещения поколений и территориальную неоднородность.

Демографический прогноз Испании основан на методе компонент. Применение этого метода соответствует следующей схеме: исходя из численности постоянного населения в определенном географическом районе и данных, наблюдаемых по каждому из основных демографических компонентов, смертности, рождаемости и миграции, ставится цель, состоящая в том, чтобы получить постоянное население на уровне более поздней даты при определенных гипотезах о будущем этих демографических явлений, определяющих их рост и их возрастную структуру.

Прогноз численности населения каждого пола, возраста и места рождения, проживающего в Испании, а также в каждом из их автономных сообществ и субъектов по состоянию на 1 января каждого года прогнозируемого периода, выполняется с помощью мультирегиональной прогнозной модели [108], которая представляет в качестве результатов не только данные о численности населения по полу и возрасту, проживающих на каждом из рассматриваемых территориальных уровней, но также прогнозируемые цифры рождаемости, смертности и миграционных движений, которые будут иметь место в каждый из моментов прогнозного периода. Все это обеспечивает необходимую согласованность между демографическими потоками и запасами, а также должную межтерриториальную согласованность.

Как отмечено выше, Национальный институт статистики Испании применяет два уровня прогнозирования: для всей страны в целом, и для субъектов государства.

Метод прогноза для субъектов реализуется через систему рекурсивных уравнений, описывающих динамику населения по возрасту, полу и миграционным характеристикам. Ниже представлены основные из них:

Уравнение (1.2): переход численности к возрасту $x+1$ с учётом смертности, эмиграции и внутрирегиональной мобильности. Позволяет определить численность населения следующего возраста на следующий год на основе текущей численности, коэффициентов смертности, эмиграции, иммиграции и межрегионального обмена. Уравнение учитывает динамическую структуру миграционных потоков и естественного движения, обеспечивая непрерывность поколенческого воспроизводства.

$$P_{h,s,x+1}^{t+1} = \frac{[1 - 0,5 * (m_{h,s,x}^t + e_{h,s,x}^t) * P_{h,s,x}^t + IM_{h,s,x}^t + \sum_{kh} e_{s,x,k,h}^t * \left(\frac{P_{k,s,x}^t + P_{k,s,x+1}^{t+1}}{2} \right) - \sum_{kh} e_{s,x,k,h}^t * \left(\frac{P_{h,s,x}^t + P_{h,s,x+1}^{t+1}}{2} \right)]}{[1 + 0,5 * (m_{h,s,x}^t + e_{h,s,x}^t)]}, \quad (1.2)$$

где $m_{h,s,x}^t$ – численность умерших за год в год t среди лиц, проживающих в субъекте h , пола s и возраста x по состоянию на 1 января года t ; $e_{h,s,x}^t$ – коэффициент эмиграции за границу в год t лиц, проживающих в субъекте h , пола s и возраста x по состоянию на 1 января года t ; $IM_{h,s,x}^t$ – представляет собой иммиграционный поток из-за границы, прибывающий в субъект h в год t лиц s -пола и возраста x по состоянию на 1 января года t ; $e_{s,x,k,h}^t$ – это коэффициент эмиграции из субъекте h в субъект k лиц s -пола и возраста x по состоянию на 1 января года t .

Уравнение (1.3) описывает расчёт численности новорождённых по полу в прогнозируемом году с учётом ожидаемого числа рождений, смертности в течение года, а также потоков иммиграции и эмиграции для младенцев. Отдельно учитывается межрегиональный обмен новорождёнными и распределение по полу.

$$P_{h,s,0}^{t+1} = N_{h,s}^t - D_{h,s,0}^t + IM_{h,s,0}^t - EM_{h,s,0}^t + \sum_k I_{k,h,s,0}^t - \sum_k O_{h,k,s,0}^t \quad (1.3)$$

где $P_{h,s,0}^{t+1}$ – численность населения пола s в возрасте 0 лет в субъекте h на начало года $t + 1$; $N_{h,s}^t$ – число родившихся пола s в субъекте h в течение года t ; $D_{h,s,0}^t$ – число умерших в возрасте 0 лет пола s в субъекте h в течение года t ; $IM_{h,s,0}^t$ и $EM_{h,s,0}^t$ – соответственно потоки иммиграции и эмиграции населения пола s в возрасте 0 лет в субъекте h в течение года t ; $I_{k,h,s,0}^t$ – входящий межрегиональный миграционный поток населения пола s в возрасте 0 лет из субъекта k в субъект h в

течение года t ; $O_{h,k,s,0}^t$ – исходящий межрегиональный миграционный поток населения пола s в возрасте 0 лет из субъекта h в субъект k в течение года t .

$$N_{h,s}^t = r_s \sum_{x=14}^{49f_{h,x}^t} \frac{P_{h,F,x}^t + P_{h,F,x+1}^{t+1}}{2} \quad (1.4)$$

где $N_{h,s}^t$ – число новорождённых пола s в субъекте h в году t ; r_s – доля новорождённых пола s среди общего числа родившихся; $f_{h,x}^t$ – коэффициент рождаемости женщин возраста x в субъекте h в году t ; $P_{h,F,x}^t$ и $P_{h,F,x+1}^{t+1}$ – численность женщин соответствующего возраста на начало и конец расчётного интервала.

Распределение новорожденных по территориям весьма актуальная задача для Санкт-Петербурга и прилегающих к нему территорий в силу желания женщин рожать в более хороших условиях и регистрировать новорожденного в Санкт-Петербурге, что может завысить численность родившихся на одной территории и занижить на другой.

Далее в компонентной схеме учитывается число умерших жителей субъекта h пола s и возраста x в течение года t . Для лиц возрастов $x = 0, 1, 2, \dots, 98$ число умерших определяется по формуле:

$$D_{h,s,x}^t = m_{h,s,x}^t \cdot \frac{P_{h,s,x}^t + P_{h,s,x+1}^{t+1}}{2} \quad (1.5)$$

где $D_{h,s,x}^t$ – число умерших в году t среди жителей субъекта h , пола s и возраста x ; $m_{h,s,x}^t$ – коэффициент смертности в году t среди жителей субъекта h , пола s и возраста x ; $P_{h,s,x}^t$ и $P_{h,s,x+1}^{t+1}$ – численность соответствующей когорты на начало и конец расчётного интервала.

Аналогичным образом определяется число эмигрировавших за границу лиц пола s , принадлежащих к поколению, достигшему возраста x на 1 января года t . Для лиц возраста $x = 0, 1, 2, \dots, 98$ число эмигрировавших рассчитывается по формуле:

$$E_{h,s,x}^t = e_{h,s,x}^t \cdot \frac{P_{h,s,x}^t + P_{h,s,x+1}^{t+1}}{2} \quad (1.6)$$

где $E_{h,s,x}^t$ – число эмигрировавших за границу в году t жителей субъекта h , пола s и возраста x ; $e_{h,s,x}^t$ – коэффициент эмиграции за границу в году t для жителей субъекта h , пола s и возраста x ; $P_{h,s,x}^t$ и $P_{h,s,x+1}^{t+1}$ – численность соответствующей когорты на начало и конец расчётного интервала.

Выполнение расчётов в соответствии с формулами (1.3) – (1.6) позволяет привести данные о рождениях, смертях и миграции к единой возрастной сетке.

Прогноз по стране сопоставляется с агрегированными результатами по отдельным территориям. Если итоговые значения расходятся, применяются корректирующие коэффициенты. Затем расчёт повторяется до тех пор, пока территориальные оценки не будут согласованы с общим прогнозом. При этом сохраняется распределение населения по полу, возрасту и месту рождения.

На заключительном этапе расчётные значения округлялись до 6 знаков после запятой. Небольшие расхождения между исходными данными и итоговыми расчётными значениями не влияют на интерпретацию результатов.

Опыт Барселоны показывает, что международные методики требуют уточнения при переходе к уровню крупного города. Для такого прогноза важны не только общие демографические показатели, но и локальные данные о миграции, маятниковых поездках и краткосрочном пребывании населения. Этот подход напрямую коррелирует с задачами прогнозирования численности населения Санкт-Петербурга, где использование принятых в официальной статистике России моделей не отражает реальной динамики мегаполиса. Опыт Барселоны наглядно показывает, что классическая триада демографического анализа – рождаемость, смертность и миграция – в условиях мегаполиса приобретает иное измерение. Их моделирование требует учёта не только общих трендов, но и территориальной мозаики, маятниковой мобильности, временных форм присутствия и воздействия городской политики на демографическое поведение горожан.

Совмещение когортно-компонентного подхода с вероятностным подходом открывает возможность выстроить сценарную рамку, чувствительную к структурным изменениям и демографическим рискам. Это уже не инструмент счёта численностей, а средство навигации в пространстве неопределённостей.

Выводы по главе 1. Мегалополис представляет собой открытую демографическую систему, в которой рождаемость и смертность дополняются сильным влиянием миграции. Официальный прогноз Росстата опирается на классический когортно-компонентный подход, однако для регионов открыто публикуется только средний вариант демографического прогноза. Для Санкт-Петербурга такая логика уже приводила к занижению прогнозной численности. В прогнозе до 2035 г. численность города оценивалась на уровне 5,501 млн человек в 2025 г. и 5,585 млн человек на начало 2036 г.²⁰, тогда как фактическая численность на начало 2024 г. уже превысила эти оценки (5,598 млн. чел.). Сходное ограничение сохраняется и в прогнозе до 2045 г., где средняя траектория не раскрывает диапазон возможных отклонений и неопределённость миграционного компонента.

Методология ООН выступает международным примером демографического прогнозирования, в котором когортно-компонентный расчёт сочетается с вероятностным подходом на национальном уровне. Однако для мегалополиса такая методология требует адаптации, поскольку городская динамика сильнее зависит от внутренней миграции, краткосрочного проживания, маятниковых перемещений. Барселона, сопоставимая с Санкт-Петербургом по масштабу и роли агломерационного центра, выступает эмпирическим примером изучения демографической динамики крупного города. Несмотря на различия в структуре населения и условиях городского развития, её опыт показывает, что демографический прогноз мегалополиса должен быть связан со стратегическим планированием и оценкой будущей нагрузки на городскую инфраструктуру.

²⁰ Предположительная численность населения Санкт-Петербурга и Ленинградской области до 2035 года : статистический бюллетень / Петростат. – Санкт-Петербург, 2020.

ГЛАВА 2. ОСОБЕННОСТИ ДЕМОГРАФИЧЕСКОЙ ДИНАМИКИ МЕГАПОЛИСА: УСЛОВИЯ И ОГРАНИЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ (НА ПРИМЕРЕ САНКТ-ПЕТЕРБУРГА)

2.1. Компонентный профиль демографической динамики

Санкт-Петербурга: баланс естественного и миграционного движения

Компонентный профиль демографической динамики любого территориального объекта формируется взаимодействием трёх процессов – рождаемости, смертности и миграции. Вклад каждого из них в траекторию численности населения различается как по масштабу, так и по степени устойчивости. Разложение общей динамики по компонентам позволяет увидеть, за счёт каких демографических механизмов город приобретает или теряет население, какие элементы воспроизводства характеризуются высокой инерционностью, а какие отличаются повышенной чувствительностью к изменениям внешней среды.

По данным Петростата, численность населения Санкт-Петербурга на 1 января 2024 сократилась на 2,28 тыс. человек²¹ по сравнению с 2023 годом, естественная убыль за 2023 год составила 12,1 тыс.²², тогда как миграционный прирост лишь 9,8 тыс. человек²³. Иными словами, миграция лишь частично компенсировала естественные потери населения и не обеспечила общий прирост. Однако уже в 2024 г. чистая миграция увеличилась до 70,5 тыс. человек, то есть более чем в 7 раз по сравнению с 2023 г. Однако столь резкий разрыв между двумя смежными годами следует интерпретировать осторожно: он может отражать не только изменение реальных миграционных потоков, но и эффект уточнения статистического учета, поскольку с августа 2024 г. Росстат совместно с МВД России перешёл от бумажной к электронной передаче первичной информации о миграции населения. Новый формат ускорил обработку данных и позволил учитывать категории мигрантов, ранее не попадавшие в статистический учет²⁴. Это подчёркивает, что миграция в условиях мегаполиса является не только

²¹ По данным ЕМИСС: Численность постоянного населения на 1 января, URL: <https://www.fedstat.ru/indicator/31557>

²² По данным ЕМИСС: Естественный прирост за год, URL: <https://www.fedstat.ru/indicator/31018>

²³ По данным ЕМИСС: Миграционный прирост, URL: <https://www.fedstat.ru/indicator/61749>

²⁴ Новости Росстата. Росстат совместно с МВД России перешёл на электронный формат передачи первичной информации о миграции населения. URL: <https://rosstat.gov.ru/folder/313/document/247959>

компенсаторным, но и наиболее изменчивым компонентом демографической динамики.

Сводная картина представлена на рисунке 7, где показана динамика естественного и миграционного прироста населения Санкт-Петербурга в 1993-2024 гг. На протяжении 1990-х и 2000-х годов естественное движение оставалось отрицательным: число умерших превышало число родившихся. Перелом произошёл в начале 2010-х годов: в 2012 г. естественный прирост впервые за рассматриваемый период стал положительным и сохранялся в 2012-2018 гг. Наиболее высокие значения естественного прироста были достигнуты в 2015-2016 гг. – 8,7 и 11,4 тыс. человек, соответственно. Однако с 2019 г. по настоящее время естественное движение вновь перешло в отрицательную зону. В 2020-2021 гг. естественная убыль резко усилилась под воздействием роста смертности вследствие пандемии COVID-19.

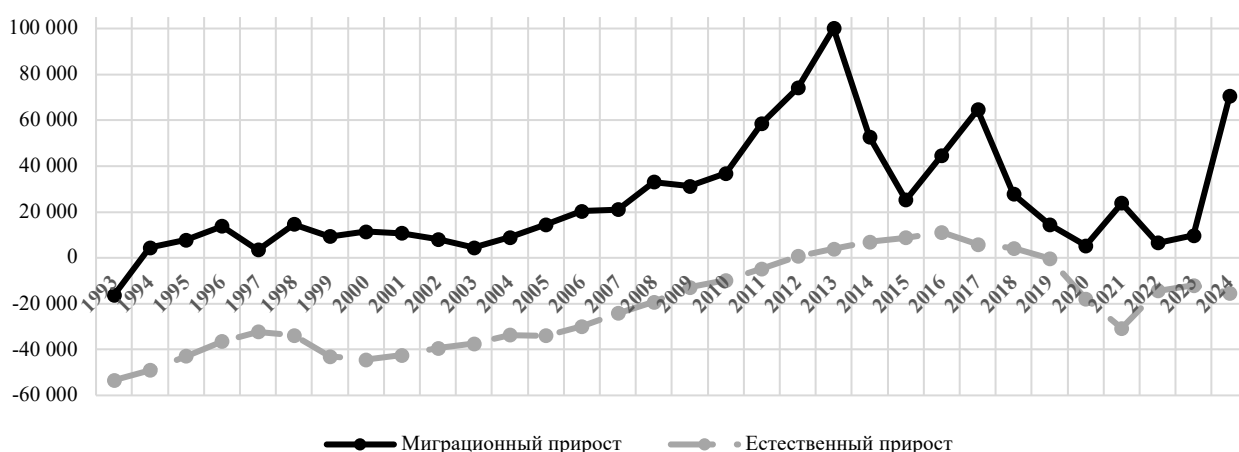


Рисунок 7 – Динамика миграционного и естественного прироста в Санкт-Петербурге, 1993 – 2022 гг.

Источник: по данным ЕМИСС «Естественный прирост за год», URL: <https://www.fedstat.ru/indicator/31018> и «Миграционный прирост», URL: <https://www.fedstat.ru/indicator/46162>

На этом фоне миграция на протяжении последних трёх десятилетий выступает ключевым источником прироста населения: миграционный прирост остаётся положительным в большинстве лет, а в отдельные периоды перекрывает масштабы естественной убыли. При исключении миграционного компонента

траектория численности населения мегаполиса приобретает нисходящий характер; фактически рост численности жителей отражает накопленный результат миграционной компенсации демографических потерь.

Тем самым уже на уровне компонентной динамики проявляется асимметрия демографического воспроизводства Санкт-Петербурга: естественное движение не обеспечивает устойчивого прироста населения, тогда как решающая роль в поддержании его численности принадлежит миграции.

2.2. Миграция как ведущий и наименее предсказуемый компонент воспроизводства мегаполиса

2.2.1. Динамика внутренней миграции

В современных мегаполисах миграция населения становится своего рода «демографическим фундаментом» устойчивого роста. Для Санкт-Петербурга она на протяжении последних трёх десятилетий выступает ключевым источником прироста населения.

Причем основной вклад вносит внутренняя миграция: с 1993 по 2023 гг. её объём вырос почти в 3 раза (относительный прирост +227%), тогда как внешний миграционный приток увеличился лишь на 32,6%²⁵. В 2023 году Санкт-Петербург принял 4,6% всех внутренних прибытий по России, а вместе с Ленинградской областью 7,0%²⁶. Эти значения свидетельствуют о значимой роли Санкт-Петербурга во внутреннем перераспределении населения и о важности агломерационного контекста при анализе миграции.

Ещё Э. Равенстейн отмечал, что люди чаще переезжают на сравнительно небольшие расстояния²⁷ [101]. Для Санкт-Петербурга это хорошо видно по составу прибывших. Значительная часть мигрантов приезжает из соседних регионов Северо-Запада, прежде всего из Ленинградской области, Республики Карелия, Новгородской и Псковской областей [53]. Такой приток объясним. Переезд из близких регионов проще, а Петербург предлагает больше возможностей для работы и учёбы. Этот «гравитационный» эффект расстояния соответствует положениям

²⁵ По данным ЕМИСС о числе прибывших по регионам России, 1990-2024 гг. URL: <https://fedstat.ru/indicator/43514>.

²⁶ По данным ЕМИСС о числе прибывших. URL: <https://fedstat.ru/indicator/43514>.

²⁷ Original (E. G. Ravenstein, 1885): "Most migrants only proceed a short distance."

гравитационного подхода, согласно которым интенсивность миграционных потоков зависит от величины взаимодействующих объектов и расстояния между ними [6, с. 84].

Внутренняя миграция Санкт-Петербурга характеризуется значительным объёмом потоков в обоих направлениях, что свидетельствует о высокой интенсивности миграционного обмена, более выраженной в Санкт-Петербурге. Для сравнительной оценки рассчитаны коэффициент прибытия и коэффициент миграционного оборота в Санкт-Петербурге и Москве в 2023 г. (табл. 2).

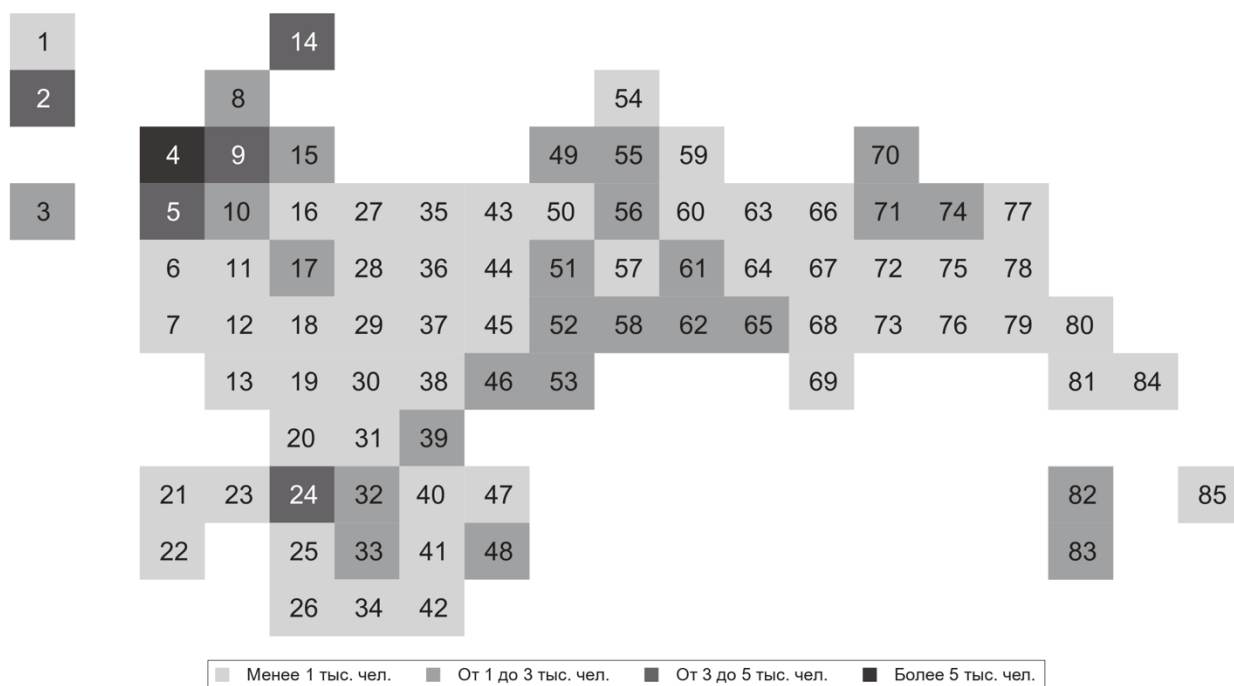
Таблица 2 – Интенсивность внутренней миграции Санкт-Петербурга и Москвы в 2023 г.

Территория	Коэффициент прибытия, ‰	Коэффициент миграционного оборота, ‰
Санкт-Петербург	28,5	58,3
Москва	22,4	43,6

Источник: Рассчитано автором на основе данных ЕМИСС: «Число прибывших» (URL: <https://fedstat.ru/indicator/43514>), «Число выбывших» (URL: <https://www.fedstat.ru/indicator/43513>), «Численность постоянного населения в среднем за год» (URL: <https://www.fedstat.ru/indicator/31556>)

Примечание: Коэффициенты рассчитываются по формулам: $K_{\text{мо}} = 1000(\Pi + B) : N$; $K_{\text{п}} = 1000(\Pi : N)$; $K_{\text{в}} = 1000(B : N)$, где Π – численность прибывших за период, B – численность выбывших за период, N – средняя численность населения региона за период.

Приток внутренних мигрантов в Санкт-Петербург складывается из двух групп потоков: из соседних и из более удалённых территорий России. Ядро донорских территорий, как уже отмечалось, составляют ближайшие регионы Северо-Запада (включая Ленинградскую область) [53, с. 45], однако существенные объёмы прибытий обеспечиваются и удалёнными регионами; в частности, устойчивый вклад фиксируется со стороны Москвы и Московской области, а также ряда субъектов вне Северо-Запада [53, с. 46] (рис. 8).



- | | | | |
|----------------------------|---|--|-----------------------------------|
| 1. г. Санкт-Петербург | 23. Республика Адыгея (Адыгея) | 45. Ульяновская область | 67. Новосибирская область |
| 2. г. Москва | 24. Краснодарский край | 46. Саратовская область | 68. Алтайский край |
| 3. Калининградская область | 25. Карачаево-Черкесская Республика | 47. Астраханская область | 69. Республика Алтай |
| 4. Ленинградская область | 26. Кабардино-Балкарская Республика | 48. Республика Дагестан | 70. Красноярский край |
| 5. Псковская область | 27. Ивановская область | 49. Архангельская область | 71. Кемеровская область – Кузбасс |
| 6. Смоленская область | 28. Владимирская область | 50. Кировская область | 72. Республика Хакасия |
| 7. Брянская область | 29. Рязанская область | 51. Республика Татарстан (Татарстан) | 73. Республика Тыва |
| 8. Республика Карелия | 30. Тамбовская область | 52. Самарская область | 74. Иркутская область |
| 9. Новгородская область | 31. Воронежская область | 53. Оренбургская область | 75. Республика Бурятия |
| 10. Тверская область | 32. Ростовская область | 54. Ненецкий автономный округ | 76. Забайкальский край |
| 11. Калужская область | 33. Ставропольский край | 55. Республика Коми | 77. Республика Саха (Якутия) |
| 12. Орловская область | 34. Республика Северная Осетия – Алания | 56. Пермский край | 78. Амурская область |
| 13. Курская область | 35. Костромская область | 57. Удмуртская Республика | 79. Еврейская автономная область |
| 14. Мурманская область | 36. Нижегородская область | 58. Республика Башкортостан | 80. Чукотский автономный округ |
| 15. Вологодская область | 37. Республика Мордовия | 59. Ямало-Ненецкий автономный округ | 81. Магаданская область |
| 16. Ярославская область | 38. Пензенская область | 60. Ханты-Мансийский автономный округ – Югра | 82. Хабаровский край |
| 17. Московская область | 39. Волгоградская область | 61. Свердловская область | 83. Приморский край |
| 18. Тульская область | 40. Республика Калмыкия | 62. Челябинская область | 84. Камчатский край |
| 19. Липецкая область | 41. Чеченская Республика | 63. Тюменская область | 85. Сахалинская область |
| 20. Белгородская область | 42. Республика Ингушетия | 64. Курганская область | |
| 21. Республика Крым | 43. Республика Марий Эл | 65. Омская область | |
| 22. г. Севастополь | 44. Чувашская Республика – Чувашия | 66. Томская область | |

Рисунок 8 – Плиточная карта «Регионы-доноры для Санкт-Петербурга в 2023 году»

Источник: составлено автором по данным ЕМИСС «Прибывшие». URL: <https://fedstat.ru/indicator/43514>

Среди удалённых регионов, представленных на карте 2023 г., выделяются Калининградская область, Краснодарский и Ставропольский края, Ростовская область, а также отдельные территории Поволжья, Урала и Сибири (Республика Татарстан, Республика Башкортостан, Республика Коми, Самарская, Свердловская, Омская, Кемеровская и Иркутская области, Красноярский край). Обращает на себя внимание присутствие в составе донорских территорий Хабаровского и Приморского краёв как наиболее удалённых восточных регионов, сохраняющих миграционные связи с Санкт-Петербургом.

Миграционное притяжение Санкт-Петербурга имеет двухконтурную структуру: ближайшее окружение формирует основной объем потока, а удаленные регионы обеспечивают его пространственную протяженность и устойчивость. Это позволяет рассматривать мегаполис не только как центр локального перераспределения населения Северо-Запада, но и как один из общероссийских узлов внутренней миграции.

2.2.2. Социально-демографический профиль внутренних и внешних мигрантов

Для оценки миграционного вклада в демографическую динамику Санкт-Петербурга недостаточно ограничиваться показателем миграционного сальдо. Не менее важно учитывать структуру миграционного притока, поскольку именно она определяет, какие группы населения пополняют численность города и каким образом миграция влияет на его возрастной, в том числе, трудоспособный и образовательный состав.

Внутренний миграционный приток в Санкт-Петербург характеризуется устойчивым преобладанием женщин. В 2023 г. женщины составляли 56% общего числа прибывших, мужчины – 43%. Данное преобладание не является равномерным по возрастной шкале (рис. 9). Такая тенденция прослеживается с 2000 г., когда число женщин среди прибывших впервые превысило число мужчин и в последующие годы, несмотря на отдельные колебания, в целом оставалось более высоким.

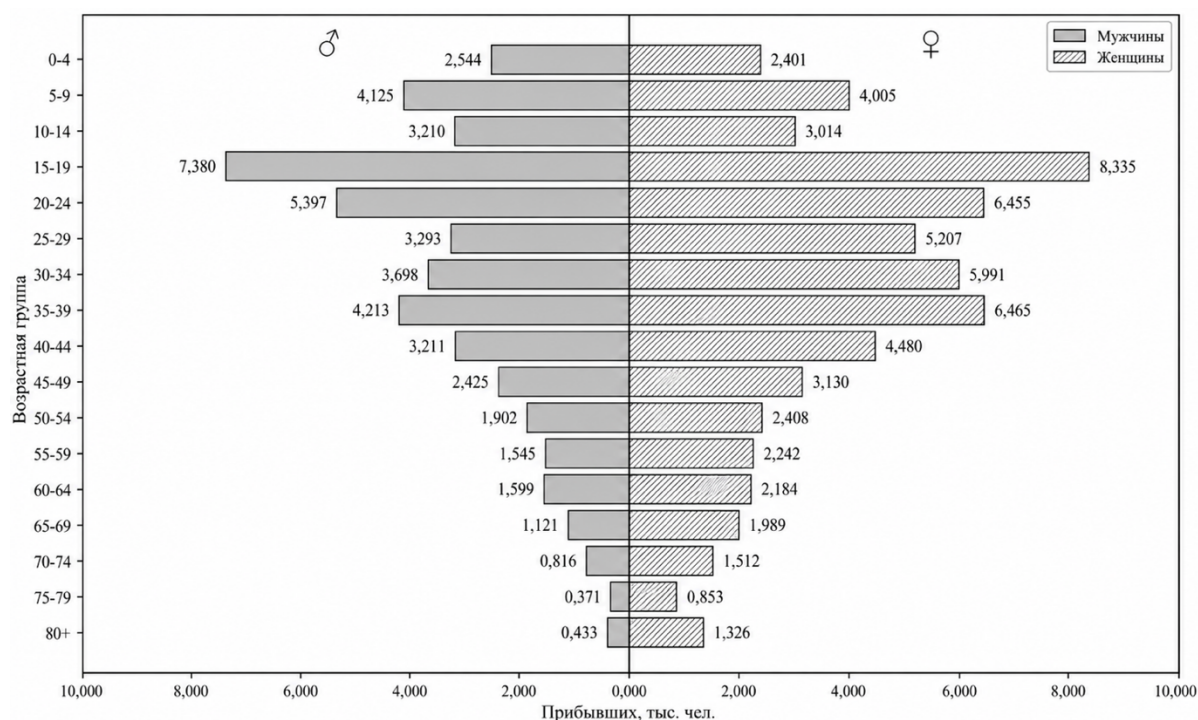


Рисунок 9 – Возрастно-половая структура внутрироссийского миграционного притока в Санкт-Петербург в 2023 г., тыс. чел.

Источник: построено автором на основе данных ЕМИСС «Число прибывших по полу, возрасту и потокам передвижения», URL: <https://fedstat.ru/indicator/58613>

В детских возрастах 0-14 лет различия между полами минимальны, в отдельных группах наблюдается даже небольшой перевес мальчиков. Начиная с возраста 15-19 лет соотношение меняется: женщины начинают преобладать практически во всех последующих возрастных группах, причём особенно заметно в наиболее мобильных трудоспособных возрастах.

Возрастной профиль внутренней миграции имеет выраженный концентрированный характер. Основное ядро притока формируют лица в возрасте 15-39 лет: на эту группу приходится свыше половины всех прибывших. Среди них практически поровну присутствуют лица в возрасте 15-24 лет (27,6 тыс. чел.) и 25-39 лет (28,9 тыс. чел.).

Интерпретация миграционного профиля молодых возрастов требует методологической осторожности. Данные о численности населения и миграции в России существенно зависят от особенностей регистрационного учёта: переехавшие в другой регион люди нередко продолжают числиться по прежнему месту жительства, а при окончании регистрации по месту пребывания

автоматически вновь учитываются по месту постоянной прописки; такие «автоматические» возвраты составляют около четверти всех учтённых переселений. При этом, «в реальности по окончании срока временной регистрации большинство мигрантов, например, молодежь, уезжающая учиться в крупные города, не возвращается к месту постоянного проживания» [37]. Следовательно, статистически наблюдаемый возрастной профиль миграции, особенно в группах, связанных с образовательной мобильностью, отражает не только реальные перемещения населения, но и специфику действующего регистрационного механизма.

Возрастная структура миграционного притока в Санкт-Петербург во многом определяется ролью города как одного из крупнейших в стране центров притяжения студенческой и молодёжной миграции [37]. «Пики притока в Центре и на Северо-Западе приходятся на студенческие возраста, что не удивительно, учитывая, что миграцией в этих округах дирижируют Москва и Санкт-Петербург, крупнейшие вузовские центры страны» [37, с.35]. Для Санкт-Петербурга это особенно показательно: «Учитывая, что население Санкт-Петербургской агломерации, и, в частности, Санкт-Петербурга, почти в 3 раза меньше, чем Московской агломерации и Москвы, интересно, что приток населения в возрасте 18 лет в агломерацию Санкт-Петербурга по абсолютным масштабам практически не отличается от притока в Московскую агломерацию. Это говорит об очень высокой привлекательности Северной столицы для студентов» [37, с.38].

Однако прибытие на учебу составляет только часть миграционного притока. В целом за 2000-2023 гг. основное миграционное пополнение Санкт-Петербурга приходилось на трудоспособные возрастные группы. Это означает, что миграция воздействует не только на общую численность населения, но и на его возрастную структуру: приток населения трудоспособных возрастов смягчает последствия старения и естественной убыли (рис. 10).

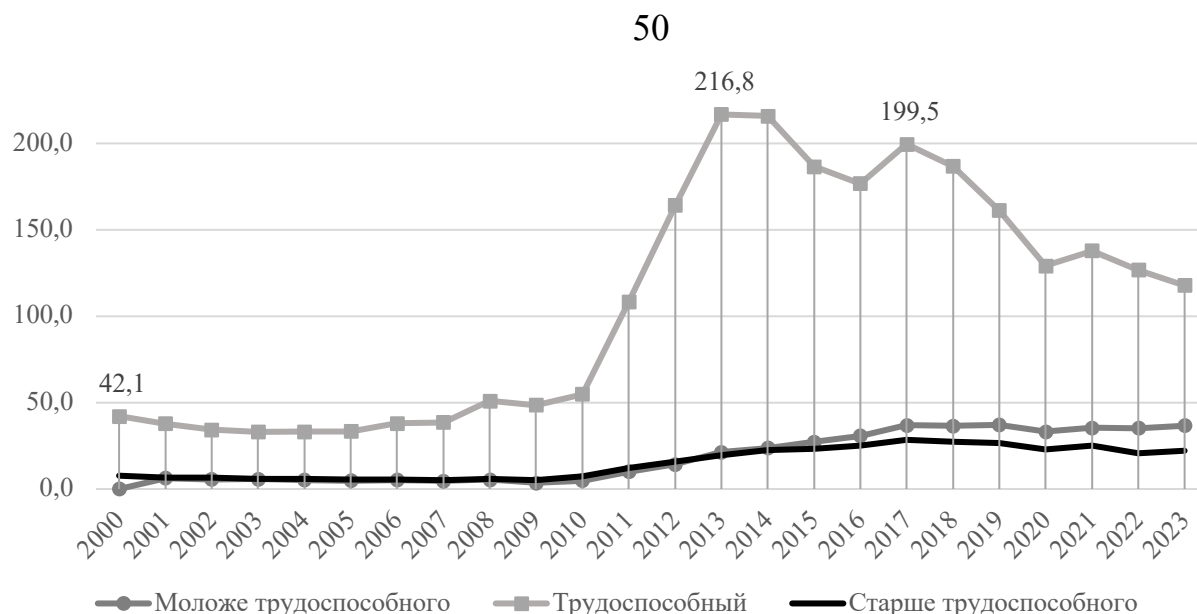


Рисунок 10 – Динамика прибывших в Санкт-Петербург по основным возрастным группам населения, 2000-2023 гг., тыс. чел.

Источник: Распределение прибывших по возрастным группам, ЕМИСС, URL: <https://www.fedstat.ru/indicator/37615>

В 2000-е гг. миграционный приток в Санкт-Петербург нарастал постепенно, в первой половине 2010-х гг. произошёл резкий подъём, особенно в трудоспособных возрастах. Этот всплеск следует рассматривать как результат совмещения реального усиления притяжения города и изменений в статистическом учёте миграции²⁸. Поэтому данные этого периода фиксируют не только рост мобильности населения, но и более полное отражение миграционных перемещений в официальных данных.

В миграционном притоке в Санкт-Петербург в 2023 г. на международную миграцию пришлось 16,9 тыс. человек, или 9,6 % общего притока. Это подтверждает, что миграционное пополнение Санкт-Петербурга связано прежде всего с его ролью центра межрегионального перераспределения населения России. Образовательная структура прибывших из других регионов включала в 2023 г. 32,9% лиц с высшим образованием, 31,9% относились к категории лиц с незаконченным высшим и средним специальным образованием, доля прибывших со средним общим образованием составляла 21,3%.

²⁸ С 2011 г. в статистический учёт долгосрочной миграции стали включаться лица, зарегистрированные по месту пребывания на срок 9 месяцев и более, а окончание срока такой регистрации стало учитываться как выбытие.

Говоря о международной миграции, стоит отметить, что в 2023 г. она перестала быть источником прироста населения. При 16,9 тыс. прибывших внешних мигрантов и 17,2 тыс. выбывших её итогом стала миграционная убыль в 255 человек. Почти весь международный приток был из стран СНГ, на которые приходилось 15,1 тыс. прибывших, или 89,1% общего объёма. Однако международный приток из стран СНГ в 2023 г. фактически свёлся к нулю. За этим фактом стоят разнонаправленные потоки из отдельных стран. Положительное сальдо сохранялось прежде всего в обмене с Таджикистаном и с Узбекистаном, в меньшей степени – с Молдовой, Туркменистаном и Азербайджаном. Напротив, обмен с Кыргызстаном, Беларусью, Казахстаном и Арменией обернулся убылью. Сальдо миграции с другими зарубежными странами оставалась и продолжает оставаться отрицательным.

Выбытие населения из Санкт-Петербурга в 2023 г., как и прибытие, было представлено в основном лицами активных трудоспособных возрастов. Наибольший отток пришёлся на возраст 20-24, 30-39 лет. Среди выбывших 25 лет и старше, как и среди прибывших, устойчиво преобладали женщины.

Отметим, что пик выбытия в группе 20-24 лет требует осторожной интерпретации. Значительная часть этих регистраций фиксирует не фактический отъезд из города, а снятие с учёта после окончания срока временной регистрации, прежде всего в общежитиях у выпускников вузов.

Причины прибытия и выбытия показывают, что миграционный обмен Санкт-Петербурга асимметричен. Приток формируется несколькими каналами закрепления в городе – воссоединение с семьей, учёбу, работу, жильё и возвращение после временного отсутствия. В выбытиях, напротив, почти половина перемещений населения 14 лет и старше приходится на возвращение после временного отсутствия; учёба и работа играют заметно меньшую роль.

Из обзора прибытия- выбытия для целей прогноза следует главный вывод: миграция Санкт-Петербурга не является однородной. Межрегиональные, внутрирегиональные и международные перемещения различаются по мотивации,

устойчивости и возрастно-половому составу, а значит, по-разному влияют на численность и структуру населения.

2.2.3. Пространственно-экономические детерминанты миграции: расширенная гравитационная модель

Пространственная организация внутренней миграции определяется не только расстоянием между регионами, но и различиями в уровне жизни и социально-экономическом развитии территорий [85]. Потоки мигрантов тяготеют к тем центрам, где выше оплата труда, шире рынок труда и возможности профессиональной и образовательной реализации. Города выступают в качестве центров экономической активности и притяжения миграции, концентрируя возможности занятости и развития [90]. Поэтому переход от описания межрегиональных миграционных связей к их факторному объяснению возможен на основе модели, учитывающей численность населения, расстояние и экономические различия между регионами.

Этой задаче отвечает расширенная гравитационная модель, в которой величина миграционного потока из региона i в Санкт-Петербург в момент времени t рассматривается как функция численности населения региона-донора, расстояния до пункта назначения и различий в социально-экономических характеристиках территорий. В основе модели лежит гравитационный принцип: интенсивность миграционного потока возрастает с увеличением численности населения региона-донора и снижается по мере роста расстояния до региона-реципиента. В применении к миграции это означает, что крупные по населению регионы обладают более значительным миграционным потенциалом, тогда как расстояние действует как фактор трения, повышающий издержки перемещения и ослабляющий устойчивость миграционных связей. В расширенной спецификации модели к этим двум опорам добавляются экономические переменные, отражающие разрыв в заработной плате и параметры занятости. В результате модель принимает вид (2.1):

$$M_{i,SP,t} = G \cdot \left(\frac{P_{i,t}^\alpha}{D_{i,SP}^\gamma} \right) \cdot e^{\delta W_{i,SP,t} + \theta E_{i,SP,t}} \quad (2.1)$$

где $M_{i,SP,t}$ – поток прибывших в Санкт-Петербург из региона i в момент времени t ; $P_{i,t}$ – численность населения региона-донора; $D_{i,SP}$ – расстояние от региона i до Санкт-Петербурга; $W_{i,SP,t}$ – показатель различий в оплате труда между регионом i и Санкт-Петербургом; $E_{i,SP,t}$ – показатель, характеризующий различия в занятости; G – константа масштаба; $\alpha, \gamma, \delta, \theta$ – параметры модели, отражающие чувствительность миграционного потока к изменению соответствующих факторов.

После логарифмирования модель принимает линейный вид, удобный для оценки параметров (2.2):

$$\ln M_{i,SP,t} = \beta_0 + \alpha \ln P_{i,t} - \gamma \ln D_{i,SP} + \delta W_{i,SP,t} + \theta E_{i,SP,t} + \varepsilon_{i,t} \quad (2.2)$$

где $M_{i,SP,t}$ обозначает число прибывших в Санкт-Петербург из региона i в году t , $P_{i,t}$ – численность населения региона-донора, $D_{i,SP}$ – расстояние от региона i до Санкт-Петербурга, $W_{i,SP,t}$ – различия в оплате труда между регионом i и Санкт-Петербургом, $E_{i,SP,t}$ – показатель занятости, β_0 – свободный член, $\varepsilon_{i,t}$ – случайная ошибка.

Эмпирическую основу составили панельные данные о миграционных потоках в Санкт-Петербург из регионов Российской Федерации за 2015-2023 гг. Выбор временного интервала 2015-2023 гг. обусловлен полнотой и актуальностью доступных официальных статистических данных Росстата о внутренней миграции. С 2015 г., после совершенствования системы миграционного и регистрационного учета в соответствии с постановлением Правительства Российской Федерации от 6 августа 2015 г. № 813, статистическая база обеспечивает более полное, достоверное и актуальное отражение перемещений населения. В качестве объясняющих переменных были включены численность населения региона-донора, расстояние до Санкт-Петербурга, различия в среднемесячной заработной плате и пропорция уровня занятости.

После построения модели были получены результаты тестирования фиксированных эффектов (табл. 3), подтвердившие статистическую значимость всех включённых переменных. Это означает, что миграционный поток в Санкт-Петербург формируется под совместным воздействием демографического

потенциала региона-донора, пространственной удалённости и экономических различий между территориями.

Таблица 3 – Результаты тестирования фиксированных эффектов смешанной модели*

Параметр модели	Степени свободы числителя	Приближенные степени свободы знаменателя**	F-статистика	Значимость
Свободный член	1	77,2	6788,9	<0,001
Население «Откуда»	1	74,1	35,5	<0,001
Расстояние, км	1	73,6	15,3	<0,001
Средняя заработная плата	1	610,5	6,2	0,01
Пропорция уровня занятости	1	610,4	6,6	0,01

* Зависимая переменная: поток миграции из региона в Санкт-Петербург.

** Дробные значения степеней свободы знаменателя являются следствием использования смешанной модели для панельных данных: при тестировании фиксированных эффектов SPSS рассчитывает приближённые эффективные степени свободы с учётом группировки наблюдений по регионам-донорам и ковариационной структуры модели. Значения 0,000 в выводе SPSS представлены как $p < 0,001$.

Источник: составлено автором по результатам построения расширенной гравитационной модели в SPSS Statistics по данным: Число прибывших. ЕМИСС. URL: <https://fedstat.ru/indicator/43514>; Расстояние между регионами Российской Федерации. URL: <https://rf.cargo-avto.ru/raschet-rasstoyaniya/> Среднемесячная номинальная начисленная заработная плата на одного работника по полному кругу организаций. ЕМИСС. URL: <https://fedstat.ru/indicator/58701> ; Пропорции уровня занятости. Сборник «Рабочая сила, занятость и безработица в России». 2024, раздел 2.2. Уровень занятости населения по субъектам Российской Федерации. URL: <https://rosstat.gov.ru/folder/210/document/13211>

Количественная оценка этих связей показывает уже не только сам факт их значимости, но и направление действия каждого фактора (табл. 4). Более населённые регионы формируют более интенсивный приток в Санкт-Петербург, пространственная удалённость этот приток ослабляет, а различия в заработной плате усиливают миграционную привлекательность мегаполиса.

Показатель занятости имеет отрицательный знак. Это означает, что более высокий уровень занятости в регионе-доноре не увеличивает, а снижает

интенсивность потока в Санкт-Петербург. Если в регионе есть рабочие места, стимул к переезду снижается. При этом различия в заработной плате сохраняют положительное влияние. Для Петербурга это показывает, что миграция определяется не только уровнем оплаты труда, но и состоянием рынка труда в регионах, откуда приезжают мигранты.

Таблица 4 – Оценки параметров модели миграционного притока в Санкт-Петербург*

Параметр модели	Оценка	Стандартная ошибка	Приближенные степени свободы знаменателя	t-статистика	Значимость	95%-й доверительный интервал	
						нижняя граница	верхняя граница
Свободный член	6,9	0,1	77,2	82,4	0	6,7	7
Население «Откуда»	0,5	0,1	74,1	5,9	0	0,3	0,7
Расстояние, км	-0,3	0,1	73,5	-3,9	0	-0,5	-0,2
Средняя заработная плата	3,2	1,3	610,5	2,5	0,01	0,8	5,7
Пропорция уровня занятости	-4,5	1,8	610,4	-2,6	0,01	-8	-1,1

* Зависимая переменная – поток прибывших из региона-донора в Санкт-Петербург. Отрицательный знак коэффициента означает обратную связь с миграционным потоком.

Полученные оценки подтверждают логику расширенной гравитационной модели: численность населения региона-донора увеличивает потенциальный объём притока, расстояние ослабляет межрегиональную связь, а различия в заработной плате усиливают притяжение Санкт-Петербурга. Отрицательный знак при показателе занятости уточняет эту зависимость: миграция ориентирована не только на более высокий доход, но и на реальные возможности включения в городской рынок труда.

Полученные оценки позволяют увидеть внутреннюю миграцию в Санкт-Петербург как результат совместного действия нескольких групп факторов – демографических, пространственных и экономических. Связь с численностью населения региона-донора имеет положительный характер: чем крупнее регион, тем выше его способность формировать миграционный приток в мегаполис. Эта

зависимость не сводится к простой пропорциональности. Увеличение численности населения на 1% сопровождается ростом притока примерно на 0,5%, что указывает на устойчивую, но неполную эластичность потока по отношению к демографической базе региона выбытия.

Расстояние продолжает сдерживать приток в Санкт-Петербург. Если расстояние увеличивается на 1%, число прибывших уменьшается на 0,3%. Переезд из близких регионов обычно дешевле и проще, поэтому именно это дает заметную часть внутреннего притока. Различия в заработной плате действуют в противоположном направлении. Более высокий уровень оплаты труда в Санкт-Петербурге усиливает приток населения, а величина коэффициента показывает заметную чувствительность миграции к разнице в доходах.

В этом отношении мегаполис притягивает не только своим масштабом, но и концентрацией экономических возможностей, которые придают миграционному движению выраженную направленность.

Вместе с тем действие экономических факторов не исчерпывается различиями в оплате труда. Обратная связь с показателем занятости вносит в эту зависимость существенное уточнение. Она показывает, что рынок труда выступает не только источником притяжения, но и средой ограничения. Само по себе преимущество в заработной плате ещё не означает обязательного усиления притока, поскольку решение о переезде определяется не только ожидаемой выгодой, но и возможностью реального трудоустройства.

2.2.4. Статистические ограничения и нерегистрируемая мобильность

Особое значение для Санкт-Петербурга имеет маятниковая трудовая мобильность в пределах агломерации, прежде всего регулярные перемещения между городом и Ленинградской областью. Моноцентричный характер Санкт-Петербургской агломерации приводит преимущественно к преобладанию центростремительных потоков рабочей силы: прилегающие районы Ленинградской области входят в зону тяготения ядра агломерации в границах Санкт-Петербурга [4, с. 87]. Оценки ежедневных объёмов маятниковых миграций между районами Ленинградской области и Санкт-Петербургом варьируют в

диапазоне от 200 до 450 тыс. человек²⁹ – это 21-47% экономически активного населения области [4]. Более осторожные расчёты, основанные на анализе пригородного пассажиропотока, дают оценку 105 тыс. человек, ежедневно въезжающих на работу в Санкт-Петербург из девяти муниципальных районов и одного городского округа Ленинградской области; эта оценка рассматривается как минимальная [4, с. 110]. Развитие промышленных производств в прилегающих к городу районах Ленинградской области формирует и обратное направление трудовых поездок: около 50 тыс. человек, или 2% экономически активного населения Санкт-Петербурга, ездят на работу из города в область. Дополнительные искажения в статистически наблюдаемую мобильность вносит сезонный дачный трафик: с мая по сентябрь на территории Ленинградской области проживает около 700 тыс. жителей Санкт-Петербурга, а в период массовых летних отпусков и школьных каникул – до 1,5-2 млн человек³⁰ [4, с. 98].

В результате статистически наблюдаемые перемещения отражают лишь часть реальной мобильности населения в пределах агломерации, оценки маятниковых трудовых миграций расходятся, носят приближенный характер: не охватывают все виды пригородных перевозок, не учитывают часть посадок в черте города и дополнительно осложняются сезонными колебаниями, связанными с эффектом дачного трафика.

Стандартный учет миграции для Санкт-Петербурга отражает не всю фактическую мобильность населения, а только ее зарегистрированную часть. Маятниковые поездки, временная трудовая занятость, краткосрочное проживание, аренда жилья без регистрации по месту жительства/ месту пребывания³¹ и сезонные перемещения создают реальную нагрузку на город, но либо не попадают в текущий учет, либо отражаются в нем косвенно [4]. Поэтому при построении прогноза

²⁹ Ежедневный приток маятниковых мигрантов в Ленинград во второй половине 1980-х годов достигал 140 тыс. человек, что составляло 11% всех занятых в экономике города [4, с. 124].

³⁰ По другим экспертным оценкам, в Ленинградской области работает около 120 тыс. жителей Санкт-Петербурга (около 4% экономически активного населения города) [4].

³¹ На основании Закона РФ от 25.06.1993 № 5242-1 и Постановления Правительства РФ от 17.07.1995 № 713 оформление регистрации по месту пребывания в жилом помещении, принадлежащем другому лицу, требует согласия собственника; следовательно, в отношениях частного найма предоставление такой регистрации выступает не как безусловная обязанность собственника, а как зависящее от его согласия условие.

используются данные о зарегистрированных прибытиях и выбытиях, характеризующие только ту часть миграционных перемещений, которая фиксируется органами статистики.

Переход Росстата с 2011 г. к учету лиц, зарегистрированных по месту пребывания на 9 месяцев и более, расширил охват долгосрочных мигрантов, но одновременно изменил объемы и структуру регистрируемых потоков [66]. Краткосрочные, циклические и нерегистрируемые перемещения при такой системе остаются за пределами полного наблюдения, а смешение источников данных, критериев учета, потоков за период и численности мигрантов на дату приводит к неверной интерпретации миграционной статистики [68]. Для прогнозирования это означает, что миграционный компонент изначально содержит более высокий уровень неопределенности, чем рождаемость и смертность, и требует вероятностного описания.

2.3. Рождаемость в условиях мегаполиса: тенденции и последствия для демографического прогноза

Рождаемость имеет особое значение для демографического прогноза, поскольку определяет численность младших групп и последующее прохождение когорт, родившихся через возрастную шкалу. В мегаполисе динамика рождаемости зависит не только от численности женщин репродуктивного возраста и интенсивности возрастной рождаемости, но и от смещения рождений к более старшим возрастам ввиду продолжительности образования, занятости, жилищных ограничений и доступности социальной инфраструктуры.

Для Санкт-Петербурга изменение возрастного профиля рождаемости является ключевым: вклад молодых групп в число рождений снижается, основная часть рождений концентрируется в более старших репродуктивных возрастах. Поэтому анализ рождаемости не может ограничиваться рассмотрением суммарного коэффициента рождаемости, а включает возрастные коэффициенты и изменение вклада отдельных возрастных групп, что необходимо для последующего когортно-компонентного прогноза [59].

Суммарный коэффициент рождаемости в Санкт-Петербурге стабильно остаётся ниже среднероссийского уровня, отражая устойчивую специфику мегаполисной фертильности. В 2023 году СКР в городе составил 1,26, тогда как в России в целом равен 1,41 (рис. 11).

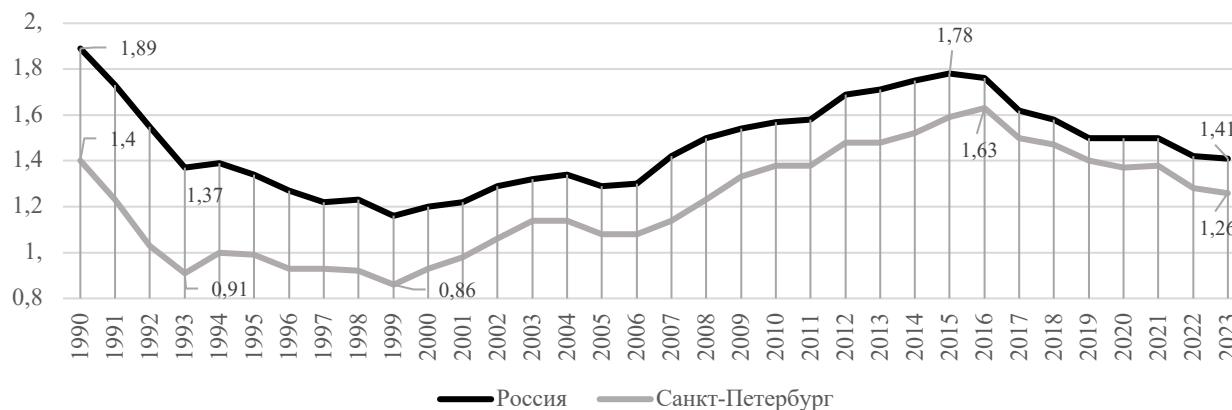


Рисунок 11 – СКР в России и Санкт-Петербурге с 1990 – 2023 гг.

Источник: составлено по данным Росстата. – URL: <https://www.fedstat.ru/indicator/31517> (дата обращения: 03.05.2024).

При этом в динамике СКР чётко прослеживаются волнообразные колебания. Подъём начала 2000-х годов во многом был связан с демографическим эхом 1980-х годов, с большими материнскими когортами. С 2007 г. дополнительное влияние на показатели рождаемости стала оказывать пронаталистская политика государства, прежде всего введение материнского капитала. Максимальные значения СКР были достигнуты в 2015-2016 гг. – в России в целом показатель приблизился к 1,78, а в Санкт-Петербурге – к 1,63 (рис. 11). Однако последующее снижение после 2016 г. показывает ограниченность инерционного действия данной меры. Существенно, что после достижения размера 453,026 тыс. руб. материнский капитал в 2016-2019 гг. не индексировался, вследствие чего его реальная покупательная способность постепенно снижалась. Поэтому дальнейший спад СКР можно рассматривать как результат наложения нескольких факторов – исчерпания демографического эха 1980-х годов, ослабления реальной стимулирующей силы материнского капитала, экономической нестабильности и последствий пандемии.

В научной литературе подобные трансформации соотносятся с концепцией второго демографического перехода, как длительного социокультурного процесса,

сопровождающегося отложенным вступлением в брак, снижением числа детей в семье, ростом индивидуализации и распространением новых форм партнёрства [95]. Санкт-Петербург, обладающий высоким уровнем урбанизации и образования, демонстрирует наиболее яркое выражение этих тенденций среди российских регионов. В Санкт-Петербурге в 2021 г. возраст матери при рождении первого ребёнка достиг 27,6 лет, что на 1,6 года выше, чем в среднем по стране³². Вместе с тем интерпретация городской рождаемости не может ограничиваться только логикой второго демографического перехода: на краткосрочную динамику могут влиять кризисы, тревожность и неопределённость жизненных перспектив. М. А. Клупт среди факторов, негативно влияющих на рождаемость, выделяет «нарастающую прекаризацию занятости «молодых взрослых» и тревожность, обусловленную неопределённостью жизненных перспектив» [26].

Другим важным индикатором является изменение возрастной структуры фертильности. Если с 1990-2000 гг. пик рождаемости приходился на возрастную группу 20-24 года, то с 2018 года он стабильно сместился в диапазон 25-29 лет³³. При этом существенно возросли показатели рождаемости в возрастах 35-39 и 40-44 лет: за последние два десятилетия – соответственно в 2,4 и 4,2 раза³⁴. Такое смещение создаёт так называемый темпо-эффект, при котором снижение текущих показателей рождаемости отчасти обусловлено отсрочкой, а не отказом от рождения детей. Темпо-эффект снижает точность линейных экстраполяций [77].

Качественные изменения в репродуктивном поведении горожан находят отражение в динамике брачности и доле внебрачных рождений – индикаторах, особенно чувствительных к ценностным и институциональным трансформациям [11]. В Санкт-Петербурге, как и в стране в целом, снижается доля ранних браков (табл. 5), сдвигается возраст заключения первого брака всё заметнее распространение устойчивых партнёрств без регистрации.

³² Молодежь СНГ – демографические штрихи к статистическому портрету 2023 года
<https://www.demoscope.ru/weekly/2024/01037/barom04.php>

³³ ЕМИСС «Возрастные коэффициенты рождаемости», URL: <https://fedstat.ru/indicator/30973>

³⁴ Рассчитано автором на основе сравнения возрастных коэффициентов рождаемости за 2003 и 2023 гг.

Таблица 5 – Возраст брачующихся в Санкт-Петербурге за 2005 и 2022 год, в процентах

Возраст жениха, лет:	2005	2022	Изменение п.п.
до 18	0,0	0,0	0,0
18-24	26,1	14,1	-12,0
25-34	45,7	48,9	3,2
35 и старше	28,2	37,0	8,8
Первый брак	68,3	67,0	-1,3
Возраст невесты, лет:	%		
до 18	0,6	0,1	-0,5
18-24	40,9	23,0	-17,9
25-34	37,8	46,3	-8,5
35 и старше	20,7	30,6	9,9
Первый брак	71,0	67,9	-3,1

Источник: Статистический бюллетень Петростата «Естественное движение населения» за 2005 и 2022 гг.

Зафиксированный характерный волнообразный сдвиг пик доли внебрачных рождений приходится на начало 2000-х годов, затем следует спад и новый подъём после 2016 года (рис. 12). Эти тенденции свидетельствуют о трансформации социальных норм родительства и постепенном размывании связи между браком и деторождением. Для мегаполиса с его высокой социальной дифференциацией и индивидуализацией жизненных траекторий это означает дополнительную неопределённость в прогнозировании рождаемости и необходимость учета нестандартных поведенческих сценариев.

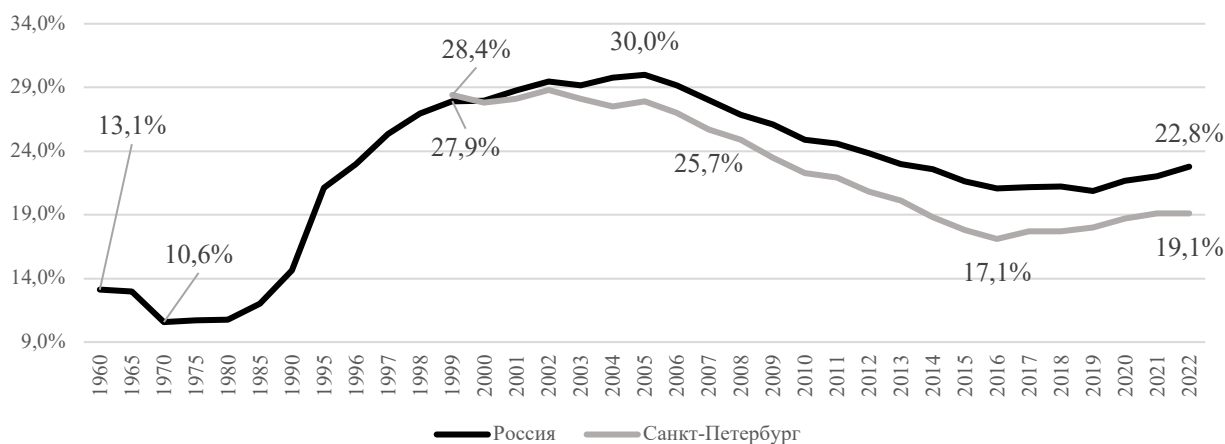


Рисунок 12 – Динамика доли внебрачных рождений в Санкт-Петербурге и России 1960 – 2022 гг.

Источник: составлено автором по данным Росстата: Социальное положение и уровень жизни населения России, раздел 2.11 «Показатели рождаемости»; Петростата: статистический бюллетень Естественное движение населения Санкт-Петербурга и Ленинградской области.

Таким образом, рождаемость в мегаполисе становится не столько следствием биологических или традиционных социальных факторов, сколько результатом институционального и ценностного контекста, варьирующегося во времени. Эти особенности делают невозможным использование жёстко заданных сценариев, построенных на фиксированных значениях СКР и возрастных коэффициентов. Модель «высокой – средней- низкой» рождаемости, применяемая в рамках традиционного компонентного подхода, оказывается методологически уязвимой в условиях высокой волатильности фертильного поведения.

При этом для всех субъектов Российской Федерации публикуется только один вариант прогноза – средний, что ограничивает возможность учёта альтернативных траекторий. Рождаемость, подверженная множеству внешних сигналов, выходит за пределы стандартных допущений и требует более гибкого инструментария – прежде всего, вероятностного.

Санкт-Петербург наглядно демонстрирует, насколько фертильность в мегаполисе может быть подвержена как институциональным сдвигам, так и поведенческим трендам, слабо поддающимся формализации. В таких условиях средняя линия прогноза теряет свою предсказательную силу. Возникает потребность в оценке диапазона возможных траекторий, каждая из которых может стать реалистичной при изменении внешнего фона. Это ставит перед демографическим анализом новую задачу: не только оценивать вероятность будущих уровней рождаемости, но и учитывать их влияние на структуру воспроизводства населения.

2.4. Смертность в мегаполисе: устойчивость показателей и значение возрастной структуры

Смертность в мегаполисе является наиболее инерционным компонентом демографической динамики населения Санкт-Петербурга, но и она не лишена внутренней сложности. В отличие от фертильности и миграции, смертность

реагирует не на колебания текущей политики или экономической конъюнктуры, а на структурные особенности населения: возраст, пол, когорты рождения, распределение по причинам смерти. Данные о смертности используются как одна из ключевых компонент когортно-компонентного прогноза, поскольку определяют вероятности дожития возрастно-половых когорт.

Возрастно-половое распределение смертности имеет выраженную неравномерность. Основная часть случаев смерти концентрируется в старших возрастах, что является общей характеристикой демографической структуры смертности. В Санкт-Петербурге устойчиво доминируют смерти в старших возрастных когортах – прежде всего среди женщин 85 лет и старше (в 2023 году численность женщин, умерших в возрасте 85 и старше составила 11,8 тыс. чел., тогда как общая численность мужчин в этой группе равна 3,1 тыс. чел.³⁵). В то же время сохраняется значительный уровень преждевременной смертности мужчин, особенно в возрасте 20-54 лет (рис. 13). Такая двойственность – «стареющая» женская смертность и «досрочная» мужская создает особенность общей смертности в мегаполисе [55].

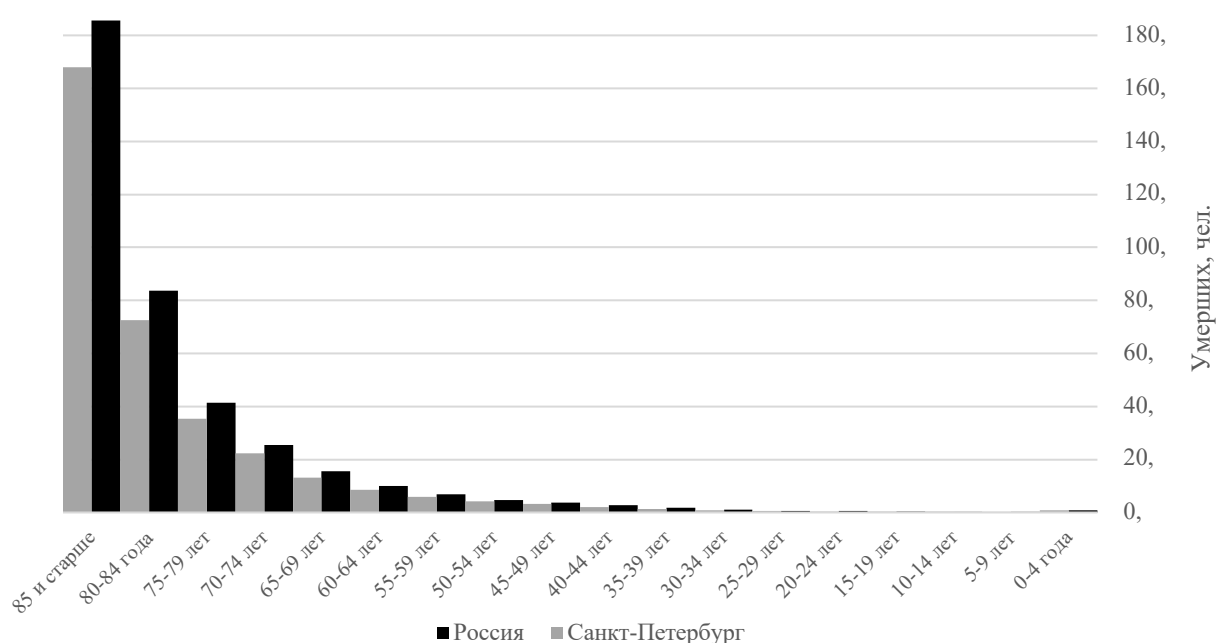


Рисунок 13 – Возрастные коэффициенты смертности в Санкт-Петербурге и России, 2022 год.

³⁵ По данным ЕМИСС «Число умерших по полу и 5-летним возрастным группам»

Благодаря институциональным ресурсам мегаполиса, включая развитую систему здравоохранения, высокий уровень образованности и доступность специализированной помощи, уровень смертности в городе стабильно ниже, чем в среднем по стране [3, 44]. В 2020-2021 гг. пандемия COVID-19 привела к росту смертности, однако в Санкт-Петербурге этот рост был менее выраженным, чем в ряде других регионов. Это позволяет говорить об относительной демографической устойчивости города в кризисный период (рис.14).

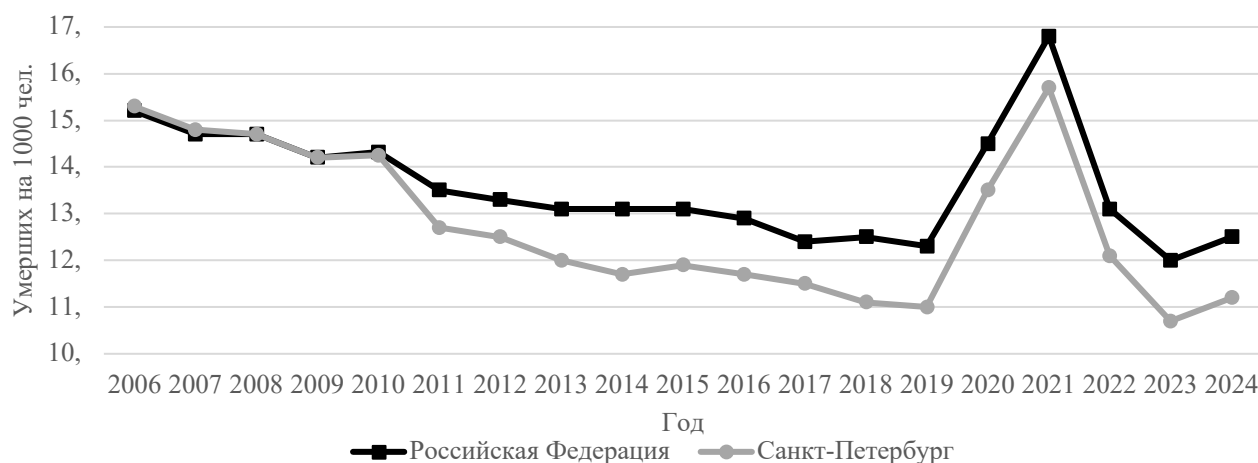


Рисунок 14 – Динамика общего коэффициента смертности в Санкт-Петербурге и России с 2006 – 2024 гг.

Источник: составлено автором по данным ЕМИСС: «Число зарегистрированных умерших в расчете на 1000 населения (оперативные данные)». URL: <https://fedstat.ru/indicator/33537>

Несмотря на общее снижение смертности, гендерный разрыв остаётся устойчивым: ожидаемая продолжительность жизни (ОПЖ) мужчин стабильно на 8 лет ниже, чем женщин³⁶. В мужской части населения смертность в трудоспособном возрасте создаёт значимую нагрузку на трудоспособное население и в целом на демографическую структуру, вызывая дефицит мужских когорт, а также высокий риск преждевременного ухода в период максимальной экономической активности. Для мегаполиса это особенно чувствительно, поскольку влияет как на численность трудоспособного населения, так и на перспективы воспроизводства.

Причинная структура смертности в мегаполисе также имеет свои особенности. Смертность от болезней системы кровообращения остаётся

³⁶ По данным ЕМИСС «Ожидаемая продолжительность жизни при рождении», URL: <https://www.fedstat.ru/indicator/31293>

доминирующей, за ней следуют ишемические болезни сердца и злокачественные новообразования³⁷. В то же время внешние причины, включая травмы, отравления и суициды, значимо сохраняются в мужской структуре смертности, особенно в младших и средних возрастах. Пандемия временно усилила смертность от инфекций, но уже к 2023 году структура смертей вернулась к обычным пропорциям.

2.5. Совокупное воздействие демографических компонентов и обоснование вероятностного подхода к прогнозированию

Если рассматривать компоненты воспроизводства совместно, то для Санкт-Петербурга становится очевидным, что их вклад в динамику численности населения неравноценен. Рождаемость и смертность в значительной степени действуют через уже сложившуюся возрастную структуру: они постепенно изменяют соотношение поколений, но сами по себе не формируют условий для устойчивого естественного прироста. Возможности рождаемости ограничиваются низкой интенсивностью рождений в молодых возрастах и концентрацией материнства в старших репродуктивных группах, тогда как влияние смертности все заметнее определяется ростом доли пожилого населения и сохраняющимися различиями в дожитии мужчин и женщин. Поэтому естественное движение выступает для города важным демографическим фоном, но не основным источником роста численности населения.

Миграция выполняет в Санкт-Петербурге компенсирующую функцию, но ее роль нельзя понимать вне исторического контекста. Послевоенное восстановление Ленинграда опиралось прежде всего на возвращение людей из эвакуации и демобилизацию, а не на свободный и устойчивый межрегиональный приток. В дальнейшем постоянное пополнение города длительное время сдерживалось паспортно-регистрационными ограничениями, что препятствовало формированию широкого миграционного ресурса. Влияние миграции всегда было волнообразно, оно усиливалось в связи с крупными стройками, особенно заметно в связи с

³⁷ Источник: ЕМИСС. – URL: <https://www.fedstat.ru/indicator/31620> (дата обращения: 16.08.2024).

созданием комплекса защитных сооружений Санкт-Петербурга от наводнений³⁸. Развитие промышленности расширяли спрос на рабочую силу, а студенческий приток увеличивал оборот населения, но не создавал значимого прироста постоянных жителей.

По оценке расширенной гравитационной модели, построенной для межрегиональных потоков в Санкт-Петербург, приток растет с увеличением численности региона-донора, снижается по мере роста расстояния и статистически значимо связан с различиями в оплате труда и параметрами занятости; для оцененной модели соответствующие коэффициенты составляют около 0,5 по размеру региона-донора и около -0,3 по расстоянию, при сохранении статистической значимости факторов заработной платы и занятости. Это означает, что миграция в Санкт-Петербург не формируется равномерным потоком со всей территории страны.

Различия в природе демографических компонентов делает методически недостаточным прогноз в виде одной фиксированной траектории. Рождаемость и смертность в большей степени подчинены возрастной инерции; миграция же быстрее реагирует на экономические различия, территориальную близость и изменения в порядке статистического учета. После реформы 2011 года в число долговременных мигрантов стали включаться лица, зарегистрированные по месту пребывания на срок девять месяцев и более, что заметно усилило учет молодых возрастов и сделало часть всплесков учебной миграции следствием регистрационного механизма, а не только завершеного переселения. Для Санкт-Петербурга это принципиально: прогноз должен учитывать не только ожидаемый уровень компонент, но и коридор их возможных отклонений. Поэтому вероятностная когортно-компонентная модель представляется не усложнением техники расчета, а адекватной формой описания будущей численности и структуры населения города.

³⁸ Комплекс защитных сооружений Санкт-Петербурга от наводнений – система дамб, водопропускных и судопропускных сооружений в Невской губе Финского залива, строительство которой началось в 1979 г. и стало одним из крупных инфраструктурных проектов позднесоветского периода.

2.6. Этапы подготовки входных данных на основе вторичной группировки

Для применения когортно-компонентной модели прогнозирования с вероятностным подходом была сформирована база данных по рождаемости, смертности, миграции и возрастно-половой структуре постоянного населения Санкт-Петербурга за 1990-2023 гг. Все показатели относятся к населению города в его административных границах и приведены к сопоставимой возрастно-половой структуре (табл. 6).

Таблица 6 – Описание базы данных для прогнозирования численности населения Санкт-Петербурга

База данных			
Рождаемость	Смертность	Миграция до 2014 года	Миграция с 2014 года
Возрастной коэффициент рождаемости	Возрастной коэффициент смертности	Прибывших и выбывших по полу и возрасту	Прибывших и выбывших по полу и возрасту
Возрастные группы:			
0-4 лет	0-4 лет	0-4 лет	0-5 лет
...	...	...	...
50-54 лет	85 и старше	80 и старше	80 и старше

Источник: данные ЕМИСС «Возрастные коэффициенты рождаемости», URL: <https://fedstat.ru/indicator/30973>, «Возрастные коэффициенты смертности», URL: <https://fedstat.ru/indicator/30974> , бюллетень Петростата «Численность и миграция населения Санкт-Петербурга»

При подготовке базы данных использовалась методика вторичной группировки: (трансформации, агрегирования и перераспределения официальных статистических данных в целях приведения их к унифицированной возрастной шкале и структурной сопоставимости. В условиях недостатка детализированных данных на всем временном интервале данных, особенно в отношении миграции, была проведена корректировка возрастных групп, усреднение пропорций и ретроспективное распределение потоков по актуальным возрастным интервалам [54].

Аналогичные принципы были использованы и при подготовке массива по смертности, где потребовалась дополнительная интерполяция и прогнозирование на основе имеющихся данных. Эти процедуры обеспечили создание целостной и

сопоставимой информационной базы, необходимой для реализации вероятностного подхода прогнозирования численности населения Санкт-Петербурга с применением когортно-компонентной модели.

2.6.1. Обоснование выбора эмпирической базы (1990-2023) и проверка чувствительности к альтернативным окнам

Эмпирическую базу для калибровки вероятностной модели составили годовые статистические ряды по Санкт-Петербургу за 1990-2023 гг.: возрастные коэффициенты рождаемости, возрастно-половые коэффициенты смертности, а также численность прибывших и выбывших по полу и возрасту. Эти показатели были приведены к согласованной возрастно-половой структуре и использованы для построения динамических рядов компонент воспроизводства населения – рождаемости, смертности и миграции. Возрастно-половая структура населения за 2023 г. далее используется как стартовое состояние когортно-компонентного прогноза.

Выбор эмпирической базы для калибровки вероятностной модели определяют не только параметры динамики, но и масштаб оцениваемой неопределённости: стохастическая дисперсия временных коэффициентов главных компонент и, следовательно, ширина прогностических интервалов выводятся из наблюдаемой изменчивости исторического периода. Поэтому исторический ряд используется не как «модель стационарного тренда», а как источник эмпирической оценки дисперсии и возможных шоков, которые должны отражаться в предиктивных интервалах вероятностного прогноза.

Принятый за базовый интервал 1990-2023 гг. обеспечивает полный охват качественно различных режимов демографической динамики. Для проверки устойчивости выбора исторического окна выполнена повторная перекалибровка модели на сокращённых периодах (1995-2023; 2000-2023; 2005-2023) с повторением РСА-разложения и переоценкой динамики временных коэффициентов. Сопоставимость результатов оценивалась как по параметрам динамики компонент, так и по калибровке прогнозных интервалов в скользящем backtesting на горизонте до 5 лет.

Уровень миграции и динамика временных коэффициентов оцениваются на полном историческом окне 1990-2023 гг., тогда как интервал 2014-2023 гг. используется только для стабилизации возрастно-половой формы базового миграционного профиля, поскольку именно для этого периода данные наиболее сопоставимы по методике учёта и детализации возрастных групп.

2.5.1. Подготовка данных о рождаемости

В данной работе использованы возрастные коэффициенты рождаемости (ASFR) для женщин в Санкт-Петербурге с 1990 по 2023 год. Данные о возрасте родившихся сгруппированы в 5-летние возрастные интервалы (15-19, 20-24, ..., 45-49 и 50-54). Коэффициенты рождаемости указаны на 1000 женщин в каждой возрастной группе в год. Например, в 1990 году ASFR составлял 34,4 (на 1000) для возраста 15-19 лет, а максимумом было значение 109,6 рождений для женщин возраста 20-24 года³⁹. Сравнение коэффициентов отражают изменение моделей деторождения с течением времени – в частности, крах в 1990-х годах и подъем в 2000-х годах. Общий коэффициент рождаемости (ОКР) в Санкт-Петербурге снижался до 6,2‰ в 1999 г., достигал 13,5‰ в 2016 г. и затем сократился до 8,6‰⁴⁰ к 2023 г. Суммарный коэффициент рождаемости рассчитывался на основе возрастных коэффициентов рождаемости как сумма ASFR по репродуктивным возрастным группам с учётом ширины возрастного интервала.

2.5.2. Подготовка данных о смертности

Для прогноза численности населения Санкт-Петербурга использованы возрастно-половые коэффициенты смертности за каждый год 1990-2023. Показатели указаны в расчете на 1000 человек соответствующей возрастной группы и пола для каждой возрастной группы (0-4, 5-9, ..., 80-84 и открытая категория 85+).

Полные ряды возрастных коэффициентов смертности по субъектам Российской Федерации, включая Санкт-Петербург, доступны в официальных

³⁹ По данным ЕМИСС «Возрастные коэффициенты рождаемости»,
URL: <https://fedstat.ru/indicator/30973>

⁴⁰ По данным ЕМИСС «Число родившихся на 1000 населения за год»,
URL: <https://fedstat.ru/indicator/31269>

публикациях Росстата только с 2011 г. Для восстановления данных о смертности за период с 1990 по 2010 годы были применены методы вторичной группировки: расчёт средних отклонений возрастных коэффициентов смертности на основе данных о городском населении в целом, что позволило точно учесть региональные особенности смертности и применить эти корректировки к данным за 1990-2010 годы, тем самым приблизительно восстановив возрастные коэффициенты смертности в Санкт-Петербурге.

Дополнительно, для восстановления данных о смертности в возрастной группе 0-4 лет в период 1990-1998 годов, использовалась экспоненциальная регрессия, так как она позволяет отражать замедляющиеся темпы снижения младенческой и детской смертности. Ввиду отсутствия о данных возрастных коэффициентах смертности для Санкт-Петербурга в период с 1990 по 2010 гг. была разработана методология корректировки на основе показателей смертности в городах России за этот период. Целью было приблизительно оценить показатели смертности в Санкт-Петербурге, учитывая различия в тенденциях смертности между Санкт-Петербургом и городским населением России с 2011 по 2023 год. Корректировка выполнялась отдельно для каждой возрастной группы и пола. Для этого рассчитывалось, насколько возрастно-половые коэффициенты смертности Санкт-Петербурга в 2011-2023 гг. в среднем отличались от соответствующих коэффициентов городского населения России. Полученные средние отклонения затем использовались как поправки при восстановлении рядов смертности Санкт-Петербурга за 1990-2010 гг.

Для количественной оценки различий в возрастных коэффициентах смертности между Санкт-Петербургом и городским населением России за 2011-2023 гг. рассчитывалось отношение коэффициента смертности Санкт-Петербурга к соответствующему коэффициенту смертности городского населения России. Расчёт выполнялся отдельно для каждого года, возрастной группы и пола по формуле (2.3):

$$R_{i,s,t} = \frac{M_{SP,i,s,t}}{M_{R,i,s,t}} \quad (2.3)$$

где $R_{i,s,t}$ – процентное отклонение возрастного коэффициента смертности в Санкт-Петербурге от среднего уровня городского населения России в году t , для возрастной группы i и пола s ; $M_{SP,i,s,t}$ – возрастной коэффициент смертности в Санкт-Петербурге; $M_{R,i,s,t}$ – возрастной коэффициент смертности городского населения России.

Значение $R_{i,s,t} > 1$ означает, что возрастной коэффициент смертности в Санкт-Петербурге выше соответствующего показателя городского населения России; значение $R_{i,s,t} < 1$ указывает на более низкий уровень смертности.

Для обобщения различий за весь период было рассчитано среднее отношение по каждой возрастно-половой группе за 2011-2023 гг. (2.4):

$$\bar{R}_{i,s} = \frac{1}{N} \sum_{t=2011}^{2023} R_{i,s,t}, \quad (2.4)$$

где $\bar{R}_{i,s}$ – среднее отношение возрастного коэффициента смертности Санкт-Петербурга к соответствующему коэффициенту смертности городского населения России за 2011-2023 гг. для возрастной группы i и пола s ; $R_{i,s,t}$ – отношение этих коэффициентов в году t ; N – число лет наблюдения, в данном случае $N = 13$; t – календарный год наблюдения, $t = 2011, \dots, 2023$.

В результате для каждой возрастно-половой группы была получена средняя относительная поправка, отражающая различие между возрастными коэффициентами смертности Санкт-Петербурга и городского населения России за 2011-2023 гг.

Полученные средние отклонения использовались для оценочной реконструкции возрастных коэффициентов смертности Санкт-Петербурга за 1990-2010 гг. на основе возрастных коэффициентов смертности городского населения России. Расчёт выполнялся для каждой возрастной группы и пола.

Для каждого года 1990-2010 гг. оценочный возрастной коэффициент смертности Санкт-Петербурга рассчитывался по формуле (2.5):

$$\hat{M}_{SP,i,s,t} = M_{R,i,s,t} \cdot \bar{R}_{i,s} \quad (2.5)$$

где $\hat{M}_{SP,i,s,t}$ – оценочный возрастной коэффициент смертности в Санкт-Петербурге для возрастной группы i , пола s и года t ; $M_{R,i,s,t}$ – возрастной коэффициент смертности городского населения России для той же возрастно-половой группы и года; $\bar{R}_{i,s}$ – среднее процентное отклонение возрастного коэффициента смертности Санкт-Петербурга от среднего уровня городского населения России за 2011-2023 гг.

Если $\bar{R}_{i,s}$ имеет положительное значение, оценочный коэффициент смертности Санкт-Петербурга оказывается выше уровня городского населения России; если отрицательное – ниже. Такая корректировка сохраняет возрастно-половую структуру смертности городского населения России за 1990-2010 гг., но масштабирует её с учётом различий между Санкт-Петербургом и городским населением страны, выявленных за 2011-2023 гг.

В результате была сформирована непрерывная серия возрастно-половых коэффициентов смертности Санкт-Петербурга за 1990-2023 гг. Для иллюстрации восстановленного участка ряда в таблице 7 приведён фрагмент данных за 1990, 1995, 2000, 2005 и 2010 гг. по отдельным возрастным группам. Полный восстановленный массив возрастно-половых коэффициентов смертности за 1990-2010 гг. представлен в приложении 5 (табл. 2).

Таблица 7 – Фрагмент восстановленного ряда возрастно-половых коэффициентов смертности Санкт-Петербурга, 1990-2010 гг., ‰

Возрастная группа, лет	Пол	1990	1995	2000	2005	2010
0-4	Девочки	3,30	2,63	2,43	1,57	1,49
0-4	Мальчики	4,38	3,44	2,78	1,88	1,59
20-24	Женщины	0,58	0,88	0,97	0,86	0,62
20-24	Мужчины	1,63	3,04	3,56	2,65	1,63
40-44	Женщины	1,99	3,36	2,91	3,44	2,64
40-44	Мужчины	5,83	11,69	10,29	11,31	7,53
60-64	Женщины	12,53	15,74	14,09	13,98	11,33
60-64	Мужчины	28,13	38,96	36,11	36,03	29,33
80-84	Женщины	93,48	102,28	99,83	89,91	84,29
80-84	Мужчины	119,07	126,94	121,63	111,95	108,9
85+	Женщины	168,38	187,78	188,27	173,96	156,33
85+	Мужчины	176,27	188,01	184,03	172,1	150,44

Источник: составлено автором на основе восстановленного входного массива возрастно-половых коэффициентов смертности, использованного при формировании базы когортно-компонентного прогноза.

Примечание: показатели приведены в расчёте на 1000 человек соответствующего пола и возраста. В таблице представлен фрагмент восстановленного участка ряда; полный массив значений за 1990-2010 гг. приведён в приложении 5 (табл. 2).

Фрагмент ряда подтверждает выраженную возрастно-половую дифференциацию смертности, поэтому в модели использовались детализированные возрастно-половые коэффициенты, а не агрегированные показатели.

Восстановление коэффициентов смертности для группы 0-4 года за 1990-1998 гг. выполнялось отдельно. Ранее описанная корректировка здесь не использовалась, поскольку для городского населения России отсутствовали сопоставимые данные для группы 0-4 года. Доступные показатели относились отдельно к младенческой смертности и смертности детей 1-4 лет, поэтому прямое сопоставление с рядом Санкт-Петербурга было затруднено. В связи с этим недостающие значения были восстановлены с помощью экспоненциальной регрессионной модели. Для 1990-1998 гг. показатели смертности детей 0-4 лет были рассчитаны отдельно. Основанием стала последующая динамика этого показателя. Экспоненциальная зависимость позволила учесть, что снижение детской смертности обычно происходит не равномерно, а с постепенным замедлением.

Первоначально был проведён анализ динамики смертности детей 0-4 лет за период 1999-2023 гг. Этот ряд показывает нелинейное снижение. В начале периода уменьшение показателя происходило быстрее, тогда как в последующие годы темпы снижения постепенно замедлялись (рис.15).

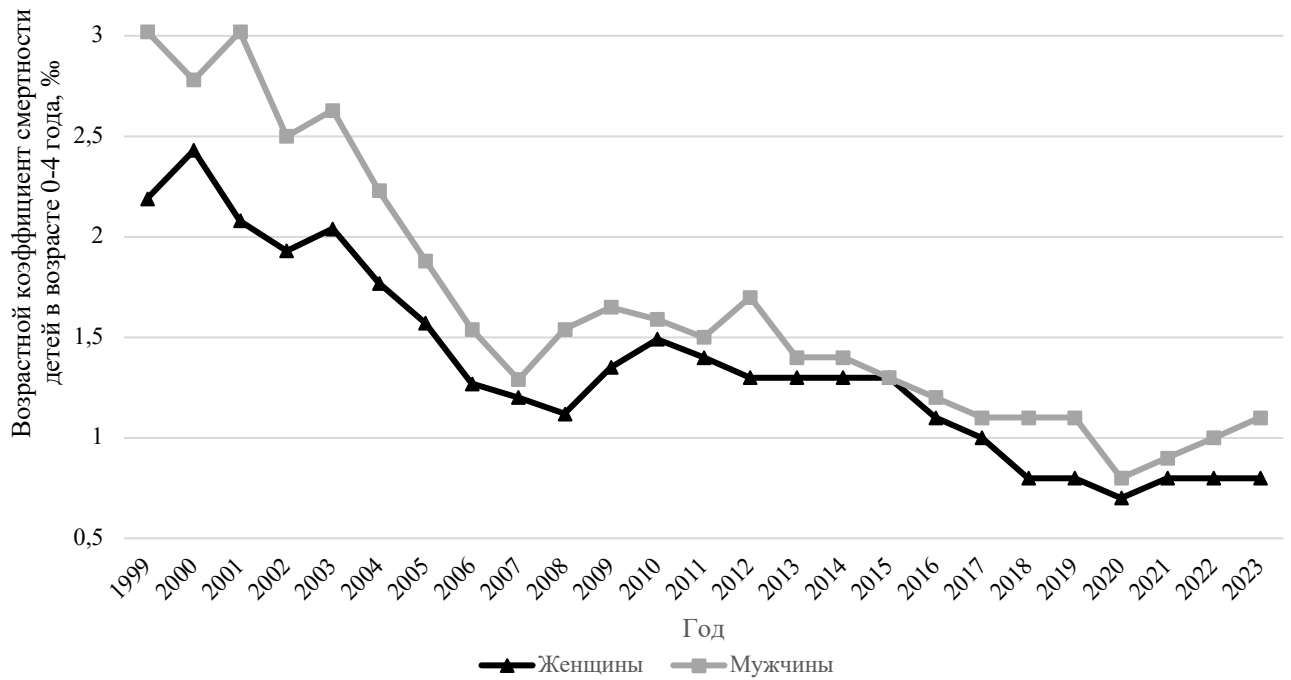


Рисунок 15 – Динамика возрастного коэффициента смертности детей в возрасте 0-4 года в Санкт-Петербурге по полу, 1999-2023 гг.

Источник: рассчитано автором по данным ЕМИСС: «Возрастные коэффициенты смертности», URL: <https://fedstat.ru/indicator/30974> ; значения за 1999-2010 гг. восстановлены автором на основе возрастных коэффициентов смертности городского населения России и средних возрастно-половых отклонений Санкт-Петербурга от городского населения России за 2011-2023 гг.

Поэтому для восстановления значений за 1990-1998 гг. была использована экспоненциальная регрессионная модель, имеющая вид (2.6):

$$m_{s,0-4,t} = e^{a_s + b_s \cdot t} \quad (2.6)$$

где $m_{s,0-4,t}$ – коэффициент смертности в возрастной группе 0-4 года для пола s в году t ; a_s и b_s – параметры экспоненциальной регрессии, оцененные отдельно для мужчин и женщин.

Для женщин модель была оценена по данным о смертности в возрасте 0-4 года в Санкт-Петербурге за последующие годы, 1999-2023 гг. Получены следующие параметры: $a = 91,364$ – свободный член модели; $b = -0,045$ – коэффициент при переменной года. Эти коэффициенты были использованы для восстановления значений за 1990-1998 гг.

С учётом полученных параметров оценочная формула имеет вид (2.7):

$$m_{F,0-4,t} = e^{91,364-0,045 \cdot \tau_t} \quad (2.7)$$

Полученные оценочные значения смертности девочек в возрастной группе 0-4 года за 1990-1998 гг. показывают снижение показателя. Например, по восстановленному ряду коэффициент смертности девочек в возрасте 0-4 года снизился с 3,3‰ в 1990 г. до 2,3‰ в 1998 г. Такая динамика соответствует общей тенденции снижения детской смертности, при которой наиболее быстрые изменения приходятся на начальные этапы, а дальнейшее уменьшение показателя происходит более медленно.

Аналогичная экспоненциальная регрессия была оценена для мальчиков по данным за 1999-2023 гг. С учётом полученных параметров оценочная формула (2.8) для мужчин имеет вид:

$$m_{M,0-4,t} = e^{97,309-0,048 \cdot \tau_t} \quad (2.8)$$

где $m_{M,0-4,t}$ – коэффициент смертности мальчиков в возрастной группе 0-4 лет в году t ; τ_t – временная переменная в шкале, использованной при оценивании регрессии.

Восстановленные значения для мальчиков также показывают снижение смертности в возрастной группе 0-4 лет, однако уровень показателя остаётся выше, чем у девочек. Смертность мальчиков снижалась с 4,4‰ в 1990 г. до 3,6‰ в 1998 г. Полученные значения использовались для формирования непрерывного входного ряда смертности за 1990-2023 гг.

Экспоненциальная модель позволила согласовать восстановленный участок ряда с фактическими данными, доступными начиная с 1999 г., и отразить общую тенденцию снижения детской смертности.

Качество аппроксимации экспоненциальных моделей, построенных отдельно для девочек и мальчиков в возрастной группе 0-4 года, подтверждается высокими значениями коэффициента детерминации: для обеих моделей R^2 превышает 0,85. Полученные ряды использовались далее как часть входного массива для когортно-компонентного прогнозирования. Входной массив возрастных коэффициентов смертности для группы 0-4 года приведён в приложении 5 (табл. 1). Он показывает,

в каком виде восстановленные значения были включены в расчётную базу когортно-компонентного прогноза: по годам, полу и возрастной группе.

2.5.3. Подготовка данных по миграции

В базу данных включены ежегодные данные о миграции за 1990-2023 годы, дезагрегированные по возрастным группам (0-4, 5-9, ..., 80-84, 85+) и полу (мужской, женский). Набор данных содержит данные о прибывших и выбывших для каждой возрастной группы и пола, из которых выводится чистая миграция. Например, в 2021 году женщины в возрасте 20-24 лет имели 9,074 тыс.чел. прибывших и 12,394 тыс.чел. выбывших (что привело к чистой миграции -3,32 тыс.чел для этой группы). Общая чистая миграция для Санкт-Петербурга значительно колебалась: в 1990 году она составляла около +17,4 тыс.чел., сократилась до +2,5 тыс.чел. к 2000 году, резко выросла до +36,8 тыс.чел. к 2010 году. В последующие годы чистая миграция колебалась весьма значительно: +23,8 тыс.чел. за 2021 год, +6,6 тыс.чел. в 2022 году и т. д., что отражает влияние пандемии и других факторов⁴¹).

При подготовке миграционных данных учитывалась неоднородность ряда за 1990-2023 гг., связанная не только с реальными изменениями миграции, но и с изменением правил текущего учета. С 2011 г. в статистику стали включаться лица, зарегистрированные по месту пребывания на срок 9 месяцев и более. Это расширило охват долгосрочных мигрантов, но изменило состав регистрируемых потоков, поскольку в ряды прибытий и выбытий стали попадать группы, ранее не учитывавшиеся как мигранты [66]. Поэтому согласование возрастных групп и восстановление недостающих фрагментов рассматривались как этап приведения наблюдаемой миграционной статистики к сопоставимому возрастно-половому формату. При интерпретации данных важно различать источник информации, критерии учета, поток за период и численность мигрантов на дату, поскольку смешение этих категорий искажает выводы о миграционной динамике [68].

⁴¹ Рассчитано автором по данным Число прибывших по полу, возрасту и потокам передвижения, ЕМИСС, URL: <https://fedstat.ru/indicator/58613>

При сборе данных о миграции, в открытом доступе были данные только за 2003-2023 гг., а за период 1990-2002 гг. не опубликованы на официальных ресурсах Росстата. В процессе прогнозирования миграционных потоков в Санкт-Петербурге для возрастных групп в период с 1990 по 2002 год был использован комбинированный подход, включающий как метод среднегодового прироста, так и модели экспоненциального и линейного роста – в зависимости от характеристик данных и особенностей динамики миграции.

На первом этапе для восстановления значений за 1990-2002 гг. использовался среднегодовой темп прироста, рассчитанный по доступному участку ряда за 2003-2010 гг. Такой подход применялся как оценочная процедура в условиях отсутствия сопоставимых прямых данных за более ранний период. Среднегодовой темп прироста определялся по формуле:

$$r = \left(\frac{X_{t+n}}{X_t} \right)^{\frac{1}{n}} - 1, \quad (2.9)$$

где X_t – значение миграционного потока в начальном году доступного интервала, в данном случае в 2003 г.; X_{t+n} – значение миграционного потока в конечном году интервала, в данном случае в 2010 г.; n – длительность интервала между наблюдениями, $n = 7$.

Применялись различные методы оценочной реконструкции в зависимости от возрастной группы и характера динамики на доступном участке наблюдений. Среднегодовой темп прироста, рассчитанный по доступному интервалу 2003-2010 гг., использовался в тех случаях, когда требовалось получить сглаженную ретроспективную оценку миграционного потока без воспроизведения годовых колебаний, которые не могут быть восстановлены по доступным данным.

В возрастных группах с более выраженными изменениями ряды прибывших и выбывших восстанавливались разными способами. Для прибывших использовалась экспоненциальная модель, поскольку в отдельных группах приток менялся неравномерно. Для выбывших применялась линейная модель, потому что их динамика была более равномерной (табл. 8). Для старших возрастов, начиная с группы 50-54 года, использовался среднегодовой темп прироста, поскольку

миграционные потоки в этих возрастах были менее интенсивными и не требовали более сложной аппроксимации.

Таблица 8 – Фрагмент рядов миграционных потоков, использованных для выбора способа восстановления данных за 1990-2002 гг., тыс. чел.

Возрастной блок, лет	Поток	2003	2007	2010	Метод
0-9	приб.	3,48	2,93	2,96	среднегод. темп прироста
0-9	выб.	2,81	2,43	2,89	среднегод. темп прироста
10-49	приб.	32,85	37,62	52,99	экспоненц. регрессия
10-49	выб.	29,28	18,6	19,67	линейная регрессия
50-54 и старше	приб.	8,44	7,74	11,18	среднегод. темп прироста
50-54 и старше	выб.	8,31	6,21	7,78	среднегод. темп прироста

Источник: рассчитано автором по данным ЕМИСС «Число прибывших по полу, возрасту и потокам передвижения», URL: <https://fedstat.ru/indicator/58613>.

Примечание: приб. – прибывшие, выб. – выбывшие; значения приведены суммарно по мужчинам и женщинам; при формировании входной базы восстановление выполнялось отдельно по полу и пятилетним возрастным группам.

Для иллюстрации восстановленного участка миграционного ряда в таблице 9 приведён фрагмент данных по прибывшим и выбывшим за отдельные годы. Поскольку миграционные потоки выражены в абсолютной численности, они могут быть агрегированы по укрупнённым возрастным группам. Более полный восстановленный фрагмент миграционных потоков за 1990-2002 гг. представлен в приложении 6 (табл. 1).

Таблица 9 – Фрагмент восстановленного ряда прибывших в Санкт-Петербург по укрупнённым возрастным группам, 1990-2002 гг., тыс. чел.

Год	Прибывшие			Выбывшие		
	0-14	15-64	65+	0-4	15-64	65+
1990	7,6	61,8	3,7	5	45,4	5,2
1995	6,8	41,9	3,8	4,7	39,1	4,5
2000	6,1	33,5	4	4,4	32,8	3,9
2002	5,9	33,5	4	4,3	30,3	3,8

Источник: составлено автором на основе восстановленного входного массива миграционных потоков, использованного при формировании базы когортно-компонентного прогноза.

Примечание: возрастные группы укрупнены для компактного представления восстановленного участка ряда; значения округлены до 0,1 тыс. чел

Полученные оценки рассматриваются не как прогноз миграции, а как способ приведения исходных данных к непрерывным возрастно-половым рядам за 1990-2023 гг. Проведенные расчеты позволили сохранить специфику миграции возрастных групп и направлением передвижения, не смешивая прибытия и выбытия в единую агрегированную величину.

1. Согласование возрастных групп миграционных данных

Корректировка потребовалась из-за изменения возрастной структуры публикуемых данных. В 2003-2013 гг. миграционные потоки представлены в одной системе возрастных групп, тогда как с 2014 г. стала использоваться более детализированная возрастная классификация (табл. 10). Для обеспечения сопоставимости всего ряда данные за период до 2014 г. методом вторичной группировки были приведены к возрастным группам, применяемым в последующем периоде. Это обеспечило единый формат входных данных для когортно-компонентной модели.

Таблица 10 – Различия в возрастных группировках мигрантов до и после 2014 года

Возрастные группы, лет	
2003-2013 гг.	с 2014 г. по настоящее время
0 – 5	0 – 4
6 – 13	5 – 9
14 – 15	10 – 14
16 – 17	15 – 19
18 – 19	

Источник: составлено автором по данным статистического бюллетеня «Численность и миграция населения Санкт-Петербурга, 2003-2023 гг.».

Для согласования возрастных групп был применён метод пропорционального распределения. На первом этапе по данным 2014-2023 гг. рассчитывались доли возрастной группы в соответствующем укрупнённом возрастном блоке (2.10).

Расчёт выполнялся отдельно для потоков прибывших и выбывших, поскольку их возрастная структура различается:

$$d_{i,t}^{(q)} = \frac{M_{i,t}^{(q)}}{M_{B,t}^{(q)}} \cdot 100\% \quad (2.10)$$

где $d_{i,t}^{(q)}$ – доля возрастной группы i в миграционном потоке типа q в году t ; $M_{i,t}^{(q)}$ – численность мигрантов возрастной группы i ; $M_{B,t}^{(q)}$ – суммарная численность мигрантов в укрупнённом возрастном блоке B , в рамках которого выполняется распределение; q – тип потока: прибытие или выбытие.

Затем годовые доли усреднялись за период 2014-2023 гг., что позволило получить устойчивую пропорцию каждой возрастной группы (2.11):

$$\bar{d}_i^q = \frac{1}{N} \cdot \sum_{t=2014}^{2023} d_{i,t}^q \quad (2.11)$$

где $\bar{d}_i^{(q)}$ – средняя доля возрастной группы i в миграционном потоке типа q за 2014-2023 гг.; n – число лет наблюдения, $n = 10$.

Полученные средние доли применялись к данным за 2003-2013 гг. для распределения укрупнённых возрастных групп по новой возрастной классификации (2.12):

$$\hat{M}_{i,y}^q = M_{B,y}^q \cdot \frac{\bar{d}_i^q}{100} \quad (2.12)$$

где $\hat{M}_{i,y}^{(q)}$ – оценочная численность мигрантов возрастной группы i в году y , $y = 2003, \dots, 2013$, для потока типа q ; $M_{B,y}^{(q)}$ – численность мигрантов укрупнённого возрастного блока B в старой классификации; $\bar{d}_i^{(q)}$ – средняя доля соответствующей возрастной группы, рассчитанная по данным 2014-2023 гг.

Такой способ не предполагает точного восстановления фактического возрастного распределения миграции в каждом году 2003-2013 гг. Его задача состоит в приведении данных разных периодов к единой возрастной структуре, необходимой для последующего когортно-компонентного расчёта. Основное допущение метода заключается в том, что средние пропорции возрастного распределения, рассчитанные по данным 2014-2023 гг., могут использоваться как приближённая структура для распределения более ранних укрупнённых возрастных групп

2. Корректировка возрастной группы 80+ на подгруппы 80-84 и 85+ в миграционных данных за 2003-2023 гг. В процессе подготовки миграционных данных за 2003-2023 гг. потребовалось обеспечить согласованность возрастных групп, используемых в разных компонентах когортно-компонентной модели. В миграционной статистике старшая возрастная группа представлена как 80+, тогда как в данных о смертности и возрастно-половой структуре населения выделяются более детализированные старшие возраста, включая группы 80-84 года и 85 лет и старше. Без дополнительной корректировки это приводило бы к несоответствию между миграционным компонентом и остальными входными рядами модели.

Для устранения этого несоответствия возрастная группа 80+ в миграционных данных была разделена на две подгруппы: 80-84 года и 85 лет и старше. Разделение выполнялось на основе структуры среднегодовой численности населения Санкт-Петербурга за 2003-2023 гг. Такой подход позволяет использовать имеющуюся информацию о возрастном составе населения старших возрастов как приближённую основу для распределения миграционного потока внутри открытой группы 80+.

Расчёт проводился для мужчин и женщин, для каждого года периода 2003-2023 гг. На первом этапе определялась доля населения 80-84 лет в общей численности населения 80 лет и старше (2.13):

$$w_{80-84,s,t} = \frac{P_{80-84,s,t}}{P_{80-84,s,t} + P_{85+,s,t}}, t = 2003, \dots, 2023 \quad (2.13)$$

где $w_{80-84,s,t}$ – доля населения 80-84 лет среди населения 80 лет и старше для пола s в году t ; $P_{80-84,s,t}$ – среднегодовая численность населения 80-84 лет в Санкт-Петербурге для пола s в году t ; $P_{85+,s,t}$ – среднегодовая численность населения 85 лет и старше для пола s в году t .

При наличии более детальной разбивки старших возрастов показатель $P_{85+,s,t}$ формировался как сумма численности групп 85-89, 90-94, 95-99 лет и более старших возрастов.

Если данных по одногодичным группам не было, то рассчитывалась доля населения 85 лет и старше как оставшаяся часть населения группы 80+ после выделения подгруппы 80-84 года (2.14):

$$w_{85+,s,t} = 1 - w_{80-84,s,t}, t = 2003, \dots, 2023 \quad (2.14)$$

На втором этапе рассчитанные доли применялись к миграционному потоку возрастной группы 80+. Распределение выполнялось для прибывших и выбывших. Для каждого типа потока использовались следующие формулы (2.15, 2.16):

$$M_{80-84,s,t}^{(q)} = M_{80+,s,t}^{(q)} \cdot w_{80-84,s,t}, t = 2003, \dots, 2023 \quad (2.15)$$

$$M_{85+,s,t}^{(q)} = M_{80+,s,t}^{(q)} \cdot w_{85+,s,t}, t = 2003, \dots, 2023 \quad (2.16)$$

где $M_{80+,s,t}^{(q)}$ – численность мигрантов в возрасте 80 лет и старше для пола s , года t и потока q – прибытия и выбытия; $M_{80-84,s,t}^{(q)}$ – оценочная численность мигрантов в возрасте 80-84 года; $M_{85+,s,t}^{(q)}$ – оценочная численность мигрантов в возрасте 85 лет и старше.

Полученные пропорции применялись к миграционным потокам возрастной группы 80+ за 2003-2023 гг. отдельно для мужчин и женщин, а также отдельно для прибывших и выбывших. В результате исходная открытая возрастная группа 80+ была разделена на две подгруппы – 80-84 года и 85 лет и старше. Это позволило привести миграционные данные к той возрастной структуре, которая используется при расчёте смертности и среднегодовой численности населения.

Выбор периода 2003-2023 гг. для расчёта пропорций связан с тем, что именно на этом интервале доступны сопоставимые данные о миграции и возрастной структуре населения, необходимые для подготовки входных рядов модели. Использование средних пропорций за весь период позволяет не привязывать распределение группы 80+ к одному отдельному году, а учитывать усреднённое соотношение между подгруппами 80-84 и 85+ в старших возрастах. Это снижает влияние годовых колебаний численности и делает распределение более устойчивым для применения ко всему миграционному ряду.

При этом такая корректировка имеет оценочный характер. Она не восстанавливает фактическое распределение мигрантов внутри группы 80+ в

каждом отдельном году, поскольку исходная миграционная статистика не содержит более детальной разбивки старших возрастов. Основное допущение состоит в том, что распределение миграционного потока внутри открытой группы 80+ приближённо соответствует возрастной структуре населения старших возрастов за тот же период.

Для иллюстрации результата разделения открытой возрастной группы 80+ в таблице 11 приведён фрагмент миграционных потоков по подгруппам 80-84 и 85+ за отдельные годы. Полный фрагмент расчётного ряда за 2003-2023 гг. представлен в приложении 6 (табл. 2).

Таблица 11 – Фрагмент разделения миграционных потоков возрастной группы 80+ на подгруппы 80-84 и 85+, тыс. чел.

Год	Пол	Прибывшие 80-84	Прибывшие 85+	Выбывшие 80-84	Выбывшие 85+
2003	женщины	0,461	0,294	0,384	0,244
2003	мужчины	0,151	0,092	0,153	0,093
2010	женщины	0,487	0,318	0,326	0,214
2010	мужчины	0,133	0,052	0,093	0,036
2014	женщины	1,054	0,899	0,811	0,692
2014	мужчины	0,294	0,188	0,203	0,129
2023	женщины	1,013	1,068	0,693	0,731
2023	мужчины	0,354	0,306	0,242	0,208

Источник: составлено автором на основе входного массива миграционных потоков, использованного при формировании базы когортно-компонентного прогноза.

Полученное разделение необходимо для согласования миграционного компонента с остальными входными данными когортно-компонентной модели. Группа 80-84 лет участвует в дальнейшем возрастном переходе, тогда как группа 85+ рассматривается как открытая верхняя возрастная категория. Обеспечение единой возрастной структуры по смертности и миграции позволило избежать несоответствий в последующем расчёте возрастно-половой динамики населения Санкт-Петербурга.

2.5.4. Стартовая численность населения

Последней составляющей когортно-компонентной модели является численность населения Санкт-Петербурга. Базовая численность населения по полу

и возрасту за 2023 г. определялась как средняя численности населения на начало 2023 г. и на начало 2024 г.:

$$\bar{P}_{a,s,2023} = \frac{P_{a,s,01.01.2023} + P_{a,s,01.01.2024}}{2} \quad (2.17)$$

где $\bar{P}_{a,s,2023}$ – среднегодовая численность населения возраста a и пола s за 2023 г.; $P_{a,s,01.01.2023}$ и $P_{a,s,01.01.2024}$ – численность населения соответствующего возраста и пола на начало 2023 и 2024 гг.

Численность представлена в 5-летних возрастных группах. В 2023 году среднегодовая численность населения Санкт-Петербурга составила 5,598 млн. чел., при этом распределение по возрасту отражает высокую представительность старших возрастных групп (58,7 тысяч женщин и 25,2 тысяч мужчин в возрасте 85+).

Для согласования возрастной структуры с прогнозной моделью старшие возрастные группы 85-89, 90-94, 95-99 лет и 100 лет и старше были объединены в открытую группу 85+ (2.18):

$$\bar{P}_{85+,s,2023} = \bar{P}_{85-89,s,2023} + \bar{P}_{90-94,s,2023} + \bar{P}_{95-99,s,2023} + \bar{P}_{100+,s,2023} \quad (2.18)$$

Объединение выполнялось отдельно для мужчин и женщин. В результате стартовая возрастно-половая структура была приведена к формату, пригодному для когортно-компонентного расчёта.

В результате вторичной группировки разнородные данные о рождаемости, смертности, миграции и численности населения за 1990-2023 гг. были приведены к единому возрастно-половому формату. Были согласованы возрастные классификации разных периодов, восстановлены недостающие значения и сформированы открытые старшие возрастные группы, необходимые для дальнейшего когортно-компонентного расчёта. Полученная база данных обеспечивает сопоставимость всех компонент модели и использовалась как входной массив для вероятностного прогноза населения Санкт-Петербурга.

Выводы по главе 2. Рождаемость, смертность и миграция отличаются по степени устойчивости и по влиянию на будущую численность населения. Рождаемость формируется в условиях мегаполиса, где рождение детей всё чаще

смещается к старшим репродуктивным возрастам, а её уровень не обеспечивает простого воспроизводства и не покрывает естественные потери населения. Влияние смертности на прогноз определяется не только возрастными коэффициентами, но и тем, какая часть населения сосредоточена в старших возрастах. Миграция остаётся наиболее изменчивым компонентом и зависит от экономической привлекательности города, межрегиональных различий, качества регистрационного учёта и перемещений, которые не полностью отражаются в официальной статистике.

Для прогноза сформирована база данных за 1990-2023 гг. Показатели рождаемости и смертности представлены в виде возрастных коэффициентов, миграция представлена через числа прибывших, выбывших и чистую миграцию, стартовая численность населения дана по полу и возрасту. Все данные приведены к единому возрастно-половому формату по пятилетним возрастным группам. Данные разных лет согласованы между собой, восстановлены недостающие фрагменты рядов, при необходимости разделены укрупнённые старшие возрастные группы, подготовлены показатели миграционного движения по полу и возрасту.

Работа с исходными данными велась в условиях ограниченности открытой официальной статистики, различий в возрастных группировках и неполноты отдельных рядов. Эти ограничения не позволяли напрямую использовать опубликованные показатели в прогнозной модели. После согласования и восстановления данных сформирован целостный массив, пригодный для вероятностного когортно-компонентного прогноза.

ГЛАВА 3. КОГОРТНО-КОМПОНЕНТНЫЙ ПРОГНОЗ ЧИСЛЕННОСТИ НАСЕЛЕНИЯ МЕГАПОЛИСА: ВЕРОЯТНОСТНЫЙ ПОДХОД

3.1. Методика вероятностного когортно-компонентного прогноза

3.1.1. Концептуальные основы

Когортно-компонентный метод описывает движение населения как результат трёх процессов – рождаемости, смертности и миграции, реализующихся в возрастно-половой структуре. Его классическая форма строится на фиксированных предположениях о будущих уровнях этих процессов. Такой подход наглядно показывает механизм демографического воспроизводства, но не отражает ту неопределённость, которая сопровождает реальные изменения показателей во времени [73].

Вероятностная модификация метода устраняет это ограничение. Вместо единственной траектории будущего формируется множество возможных вариантов, каждый из которых отражает различное сочетание исходных условий и случайных отклонений [46]. Рождаемость, смертность и миграция рассматриваются как процессы с естественными колебаниями, а не как заданные константы. Исследования показывают, что прогнозы, основанные на жёстких сценариях, не отражают неопределённость и, следовательно, риск отклонений от ожидаемой траектории [48].

Вероятностный когортно-компонентный подход позволяет оценивать не только ожидаемое значение численности населения, но и диапазон её возможных изменений. Для каждого года прогнозного периода формируется распределение вероятных исходов, на основе которого определяются медианные значения и предиктивные интервалы. Такая форма прогноза отражает вероятные границы неопределённости и делает демографические расчёты более устойчивыми и интерпретируемыми [81].

Стохастический подход приобретает особое значение на региональном и городском уровне, где неопределённость демографической динамики проявляется сильнее, чем в национальных оценках [110]. Главным источником колебаний выступает миграция, характеризующаяся высокой изменчивостью во времени и

способностью менять направление потоков в пределах коротких периодов. В исследовании Н. Кейльмана (Niko Keilman) показано, что среднеквадратическая ошибка прогноза для норвежских регионов в два-три раза выше, чем для страны в целом, главным образом вследствие флуктуаций чистой миграции [91]. М. Камерон и Ж. Пут (Cameron Michael, Poot Jacques) показали, что для субрегионов Уайката в Новой Зеландии разброс ошибок миграции превышает вариацию рождаемости и смертности в пять-шесть раз [81]. В аналогичных исследованиях Т. Уилсона (Tom Wilson) для Австралии миграционные колебания также объясняют большую часть совокупной ошибки прогнозов численности населения на уровне округов [109].

Как уже отмечалось, классическая когортно-компонентная модель описывает численность населения как результат взаимодействия трёх процессов – рождаемости, смертности и миграции. Для возрастной группы a и года t баланс записывается в стандартной форме:

$$P_{a,t+1} = P_{a-1,t} \cdot S_{a-1,t} + M_{a,t}, \quad (3.1)$$

где $P_{a,t}$ – численность населения данного возраста в начале периода, $S_{a-1,t}$ – вероятность дожития, то есть перехода из группы $(a - 1)$ в возрастную группу a , $M_{a,t}$ – сальдо миграции возрастной группы a за год t .

Для возрастной группы 0 лет (новорожденные) численность определяется рождаемостью:

$$P_{0,t+1} = \sum_{a \in A_f} P_{a,t}^{(f)} \cdot F_{a,t}, \quad (3.2)$$

где $P_{a,t}^{(f)}$ – численность женщин детородного возраста, $F_{a,t}$ – возрастной коэффициент рождаемости в год t .

В стохастическом (вероятностном) варианте когортно-компонентная модель сохраняет общую структуру демографического баланса, но допускает, что каждая её составляющая – рождаемость, смертность и миграция – может отклоняться от средних траекторий [71]. Иными словами, каждое уравнение модели дополняется стохастическими компонентами, отражающими неопределённость демографических процессов.

В общей форме вероятностный демографический баланс записывается как (3.3):

$$P_{a,t+1} = (P_{a-1,t} \cdot S_{a-1,t})(1 + \varepsilon_{S,a,t}) + (M_{a,t} + \varepsilon_{M,a,t}), \quad (3.3)$$

где $\varepsilon_{S,a,t}$ и $\varepsilon_{M,a,t}$ - случайные отклонения, характеризующие вариации смертности и миграции относительно ожидаемых значений.

Для возрастной группы 0 лет (новорождённые) также вводится стохастическая компонента рождаемости (3.4):

$$P_{0,t+1} = \sum_{a \in A_f} P_{a,t}^{(f)} F_{a,t} (1 + \varepsilon_{F,a,t}), \quad (3.4)$$

где $\varepsilon_{F,a,t}$ описывает флуктуации рождаемости у женщин возраста a под воздействием экономических, социальных и поведенческих факторов.

На практике стохастические компоненты моделируются с использованием временных рядов, что обеспечивает реалистичное воспроизведение исторической изменчивости показателей.

Формализованное описание стохастических компонент для каждой демографической составляющей имеет следующий вид. Для смертности вероятность дожития $S_{a,t}$ описывается экспоненциальной функцией стохастического коэффициента смертности:

$$S_{a,t} = \exp[-(m_{a,t} + \varepsilon_{M,t})], \quad (3.5)$$

где $m_{a,t}$ - средний прогнозный уровень смертности, $\varepsilon_{M,t}$ - случайная компонента, отражающая колебания относительно ожидаемой траектории.

Для миграции сначала задаётся возрастная ставка чистой миграции $r_{a,t}$, к которой добавляется случайное отклонение $\varepsilon_{R,t}$. Затем эта ставка умножается на численность населения соответствующей возрастной группы $P_{a,t}$, что позволяет получить абсолютную величину миграционного сальдо (3.6):

$$M_{a,t} = P_{a,t} \cdot (r_{a,t} + \varepsilon_{R,t}), \quad (3.6)$$

где $M_{a,t}$ - абсолютное миграционное сальдо возрастной группы a в году t ; $P_{a,t}$ - численность населения возрастной группы a в году t ; $r_{a,t}$ - возрастная ставка

чистой миграции; $\varepsilon_{R,t}$ – случайное отклонение, отражающее изменение интенсивности миграционных потоков [96].

Формула (3.6) представляет стохастическую модификацию компонента классической когортно-компонентной модели. Для регионального прогноза, важно в каких возрастах и по каким группам формируется миграционный приток или отток. Миграция редко распределяется равномерно по населению, поэтому в модели используются ставки по полу и возрасту (rate-based approach) [81]. Такой подход применён М. Камерон и Ж. Путом для региона Уайкато в Новой Зеландии, где чистая миграция рассматривалась через территориальные возрастно-половые ставки [81]. Это позволяет рассчитывать миграцию с учётом численности соответствующих возрастно-половых групп и использовать полученные ставки для построения стохастических прогнозных траекторий.

Использование ставок миграции в настоящем исследовании представляет собой нововведение в рамках стохастического когортно-компонентного подхода на субнациональном уровне России. В большинстве вероятностных реализаций [46, 81] стохастические компоненты добавляются к чистому сальдо миграции без перехода к возрастно-половым ставкам.

Случайные отклонения вводятся аддитивно, а не мультипликативно, поскольку исторические колебания коэффициентов рождаемости, смертности и миграции проявляются в отклонениях от средних уровней, а не в пропорциональных изменениях их масштабов. Аддитивная форма сохраняет интерпретацию модели в линейных терминах и позволяет использовать методы анализа временных рядов (например, ARIMA) для моделирования динамики случайных составляющих. При этом сохраняется возможность обратного преобразования коэффициентов в вероятностные интервалы исходных показателей, что делает модель сопоставимой с эмпирическими данными.

Вероятностный когортно-компонентный подход опирается на три методических положения. Первое – трактовка рождаемости, смертности и миграции как случайных процессов с устойчивыми возрастно-половыми профилями. Второе – формирование множества стохастических траекторий

численности населения, отражающих вариабельность компонент. Третье – представление результатов расчётов в виде медианных оценок и калиброванных предиктивных интервалов, интерпретируемых как вероятностные характеристики будущих исходов.

В современных исследованиях эти принципы получают дальнейшее развитие. Вероятностное моделирование всё чаще используется не только для отражения случайности демографических траекторий, но и для учёта внутренней неоднородности населения. Переход от усреднённых коэффициентов к дифференцированным профилям рождаемости, смертности и миграции позволяет описывать не только общий объём, но и структуру будущего населения. Опыт стран с развитой статистической базой показывает, что различия по уровню образования и социальным характеристикам формируют устойчиво отличающиеся демографические режимы. Игнорирование этих различий снижает точность не только средней оценки численности, но и пределов неопределённости. Для крупных городов это особенно существенно. Вероятностная модель, способная учесть такую избирательность миграции и естественного движения, даёт более реалистичное описание будущей динамики численности и состава населения [48, 117].

Важной задачей становится также обеспечение достоверности предиктивных интервалов прогноза. В современной практике используется эмпирическая калибровка интервалов, когда их ширина определяется не экспертно, а на основе фактических ошибок прошлых прогнозов. Ретроспективная проверка (backtesting) позволяет оценить, как часто реальные значения попадали бы в рассчитанный диапазон. Если покрытие оказывается ниже заданного уровня, интервалы дополнительно калибруются. В субнациональных прогнозах, в том числе для округов США, такие процедуры применяются для сопоставления результатов между территориями и количественной оценки надёжности прогноза [110, 117].

Вероятностный прогноз показывает не только центральную траекторию, но и возможный разброс будущих значений. Для демографических расчётов это особенно важно, поскольку неопределённость возникает по всем основным

компонентам воспроизводства населения, включая рождаемость, смертность, миграцию и итоговую численность. Поэтому задача модели состоит не только в получении ожидаемой траектории, но и в оценке диапазона, в пределах которого могут находиться будущие значения [54]. Идея о необходимости явного представления неопределённости восходит к классическим работам по пределам демографического прогнозирования, где подчёркивается, что долгосрочный прогноз не может рассматриваться как точное предсказание единственной будущей траектории [92]. Такой подход делает прогноз не предсказанием, а инструментом анализа демографических рисков и реальных границ возможного будущего. В отечественной литературе аналогичная логика проявляется в долгосрочных многовариантных прогнозах, где будущая численность населения рассматривается через комбинации гипотез рождаемости, смертности и миграции, а результаты интерпретируются как диапазон возможных демографических траекторий [69].

3.1.2. Стохастическое моделирование компонент (PCA-AR(1))

Стохастическое моделирование компонент реализуется на основе представления рождаемости, смертности и миграции как возрастно-специфических временных процессов [89]. Для каждой компоненты формируются последовательности коэффициентов $x_{a,t}$, где a обозначает возрастную группу, а t – календарный год наблюдения. Таким образом, каждый демографический процесс представляется в виде матрицы наблюдений:

$$\mathbf{X} = \begin{bmatrix} x_{a_1,t_1} & x_{a_1,t_2} & \cdots & x_{a_1,t_T} \\ x_{a_2,t_1} & x_{a_2,t_2} & \cdots & x_{a_2,t_T} \\ \vdots & \vdots & \ddots & \vdots \\ x_{a_A,t_1} & x_{a_A,t_2} & \cdots & x_{a_A,t_T} \end{bmatrix}, \quad (3.7)$$

Строки отражают возрастные профили показателя, а столбцы – годы исторического периода. Такая форма данных позволяет рассматривать динамику по возрасту и времени в едином аналитическом пространстве, что является предпосылкой для последующего применения методов главных компонент и авторегрессионного моделирования.

В исследовании используются три массива: возрастные коэффициенты рождаемости женщин $ASFR_{a,t}$ (в ‰), возрастно-половые коэффициенты

смертности $ASMR_{a,t}^{(s)}$ (в ‰) и чистые миграционные потоки $M_{a,t}^{(s)}$ (в абсолютных значениях), где $s \in \{F, M\}$. Исходные данные за 1990-2023 гг. преобразуются в согласованный панельный формат «год × возраст × пол», обеспечивающий совместимость структур компонент. В программной реализации согласование исходных таблиц выполнялось с использованием операции свёртки данных по году, возрастной группе и полу. Пример соответствующего фрагмента кода приведён в листинге 1:

Листинг 1 – Формирование возрастно-временных матриц демографических компонент

```
# Преобразование исходных таблиц в матричный формат для прогноза

# Смертность: возрастно-половые коэффициенты смертности, Year × (Sex, Age_Group)
mort_w = wide(mort, "Mortality_Rate", sex=True)

# Рождаемость: возрастные коэффициенты рождаемости женщин, Year × Age_Group
fert_f = fert.query("Sex == 'F'").copy()
fert_w = wide(fert_f, "Fertility_Rate", sex=False)

# Миграция: прибытия, выбытия и чистая миграция по полу и возрасту, Year × (Sex, Age_Group)
mig_in_w = wide(mig, "Migration_Inflow", sex=True)
mig_out_w = wide(mig, "Migration_Outflow", sex=True)
mig_net_w = mig_in_w - mig_out_w
```

Примечание. Полная версия программного кода представлена в приложении 4.

Такая форма данных позволяет представить возрастные профили демографических показателей как векторы наблюдений, индексированные по году, и использовать их в последующем стохастическом моделировании.

Базовая возрастно-половая структура населения Санкт-Петербурга за 2023 год фиксируется как исходное состояние прогноза ($P_{a,2023}^{(s)}$). В программной реализации это выполнялось следующим образом (листинг 2):

Листинг 2 – Формирование сводной таблицы численности населения по полу и возрасту (базовый год 2023)

```
# Формирование сводной таблицы численности населения по полу и возрасту
# для базового года прогноза

pop2023 = pop.query("Year == @START_YEAR").copy()

pop_wide = pop2023.pivot_table(
    index="Age_Group",
    columns="Sex",
```

```

        values="Population",
        aggfunc="sum",
        observed=False
    ).fillna(0)

pop_wide = pop_wide.reindex(
    index=ordered_age_cat(pop2023["Age_Group"]).categories
)

pop_wide = pop_wide.rename_axis(
    index="Age_Group",
    columns="Sex"
)

```

Примечание. Полная версия программного кода представлена в приложении 4.

Эта структура выступает точкой отсчёта когортно-компонентного расчёта: в каждом стохастическом сценарии она развивается во времени под воздействием прогнозных коэффициентов рождаемости, смертности и миграции, формируя траекторию численности населения вплоть до 2045 года. Период прогнозирования ограничен 2045 г., поскольку последний доступный официальный демографический прогноз Росстата для Санкт-Петербурга опубликован до 2045 г. Такой выбор обеспечивает сопоставимость авторских расчетов с официальной прогнозной траекторией.

Перед построением стохастической модели производится верификация уровней демографических компонент в стартовом году.

Для рождаемости оценивается суммарный коэффициент рождаемости (TFR) по формуле:

$$TFR_t = \sum_{a \in A_f} \frac{F_{a,t} \cdot 5}{1000}, \quad (3.8)$$

Расчет реализуется программно (Листинг 3).

Листинг 3 – Функция вычисления суммарного коэффициента рождаемости (TFR) по возрастным группам

```

# Функция расчёта суммарного коэффициента рождаемости
# на основе возрастных коэффициентов рождаемости

def compute_TFR_row(asfr_row, group_labels):
    # asfr_row: значения ASFR в промилле для выбранных возрастных групп
    width_years = 5

```

```

vals = asfr_row.reindex(group_labels).astype(float).fillna(0.0)
tfr = (vals.sum() * width_years) / 1000.0
return tfr

```

Примечание. Полная версия программного кода представлена в приложении 4.

Аналогично проверялись средние уровни смертности по полу и чистое сальдо миграции. Эти проверки использовались для контроля исходной базы перед запуском стохастического моделирования и позволяли исключить очевидные аномальные значения.

Для стохастического моделирования компонент применялось сочетание РСА и авторегрессионного моделирования. Использование компонентного представления возрастных рядов согласуется с подходом Р. Хайндмана и Х. Бута к вероятностному прогнозированию смертности, рождаемости и миграции [89]. Динамика временных коэффициентов компонент задавалась моделью $AR(1)$, эквивалентной $ARIMA(1,0,0)$ [86].

Для стохастического моделирования компонент применялось сочетание РСА и авторегрессионного моделирования [80, 89]. Метод главных компонент использовался для сжатого описания возрастных рядов, а динамика временных коэффициентов компонент задавалась моделью $AR(1)$, эквивалентной $ARIMA(1,0,0)$. Это позволило сохранить возрастные различия показателей и задать их дальнейшую динамику в вероятностной форме.

Перед проведением анализа значения коэффициентов преобразуются в логарифмическую шкалу, что устраняет масштабные различия между возрастными группами и предотвращает появление отрицательных значений при обратном преобразовании. Пусть X – матрица логарифмированных коэффициентов размером $T \times A$, где T – число лет наблюдения, A – количество возрастных групп. После вычисления средних значений μ каждый вектор возрастных коэффициентов представляется в виде линейной комбинации нескольких главных компонент:

$$\ln X_t = \mu + \sum_{k=1}^K z_{k,t} v_k + \varepsilon_t, \quad (3.9)$$

где v_k – k -я главная компонента (возрастной паттерн), $z_{k,t}$ – её временной коэффициент, ε_t – остаток. На практике двух компонент $K=2$ достаточно, чтобы описать основную часть вариации возрастных профилей. Эмпирическая проверка этого приведена ниже.

Для проверки достаточности двухкомпонентного представления была рассчитана доля дисперсии возрастных профилей, объясняемая первой и второй главными компонентами. Расчёты выполнены по демографическим компонентам и полу: для смертности и миграции – по мужчинам и женщинам, для рождаемости – по женщинам репродуктивных возрастов (табл. 12).

Таблица 12 – Доля объяснённой дисперсии первых двух главных компонент, %

Компонента	PC1, %	PC2, %	PC1 + PC2, %
Смертность, женщины	73,88	12,79	86,67
Смертность, мужчины	77,63	11,71	89,34
Миграция, женщины	74,05	10,36	84,41
Миграция, мужчины	66,83	14,77	81,6
Рождаемость	87,95	10,79	98,74

Источник: расчёты автора.

Первые две главные компоненты объясняют от 81,60 до 98,74% вариации возрастных профилей. Наибольшая доля объяснённой дисперсии получена для рождаемости, что связано с более устойчивой формой возрастного профиля. Для миграции совокупная доля ниже, особенно у мужчин, что отражает более выраженную изменчивость возрастной структуры миграционных потоков. Тем не менее во всех случаях двухкомпонентное представление сохраняет основную вариацию возрастных профилей.

Временная динамика коэффициентов $z_{k,t}$ моделируется авторегрессионным процессом первого порядка:

$$z_{k,t} = \varphi_k z_{k,t-1} + \eta_{k,t}, \quad (3.10)$$

где φ_k – параметр авторегрессии, $\eta_{k,t} \sim N(0, \sigma_k^2)$ – случайная ошибка. Модель AR(1) позволяет учитывать инерционность демографических изменений, не вводя искусственных трендов.

При выборе конкретной спецификации авторегрессионной модели была проведена диагностика альтернативных ARIMA-процессов для временных

коэффициентов главных компонент. Сравнивались модели различного порядка на основе информационных критериев AIC и BIC для нескольких вариантов исторического окна наблюдений (1990-2023, 1995-2023, 2000-2023 и 2005-2023 гг.; (прил. 1, табл. 1). Результаты показали, что «лучшая» модель по критериям информационной оптимальности зависит от выбранного интервала наблюдений: при изменении длины временного ряда оптимальная спецификация ARIMA может изменяться. В этих условиях использование более сложных ARIMA-спецификаций не обеспечивает устойчивого преимущества и увеличивает риск переобучения модели.

По этой причине в работе используется парсимоничная⁴² спецификация AR(1)⁴³ [86], соответствующая авторегрессионному процессу первого порядка. Такая модель отражает инерционную природу демографической динамики и одновременно обеспечивает стабильность параметров при различных вариантах исторического окна. Для задач вероятностного прогнозирования это является принципиальным преимуществом, поскольку позволяет воспроизводимо генерировать стохастические траектории компонент воспроизводства населения в рамках процедуры Монте-Карло.

Для каждого $z_{k,t}$ вычисляются прогнозные значения $\hat{z}_{k,t}$ и стохастические траектории, полученные добавлением нормально распределённого шума $\eta_{k,t}^{(d)}$. Повторяя симуляции $d = 1, \dots, N_{sims}$ (в исследовании $N_{sims} = 10\,000$), формируется ансамбль возможных траекторий компонент, что обеспечивает вероятностное представление будущих изменений.

Выбор эмпирической базы для калибровки вероятностной модели определяют не только параметры динамики, но и масштаб оцениваемой неопределённости: дисперсия временных коэффициентов главных компонент и,

⁴² Парсимония (от лат. Parsimonia – экономность) – принцип построения статистической модели, согласно которому предпочтение отдаётся более простой модели с меньшим числом параметров при сопоставимом качестве аппроксимации данных. В контексте моделирования временных рядов это означает выбор минимально необходимого порядка авторегрессионной или ARIMA-модели, позволяющего адекватно воспроизводить наблюдаемую динамику и избегать переобучения модели.

⁴³ Orig.: «A similar model with an AR(1) term added was marginally superior, but we preferred the (0, 1, 0) model on grounds of parsimony» [104].

следовательно, ширина прогностических интервалов выводятся из наблюдаемой изменчивости исторического периода. Поэтому исторический ряд используется не как «модель стационарного тренда», а как источник эмпирической оценки дисперсии и возможных шоков, которые должны отражаться в предиктивных интервалах вероятностного прогноза.

Базовый интервал охватывает 1990-2023 гг., что обеспечивает присутствие качественно различных режимов демографической динамики. Такой подход соответствует общему требованию демографического прогнозирования использовать достаточно длинную базу наблюдений, поскольку сокращение исторического окна может снижать устойчивость параметров модели [93, с. 659]. Для проверки устойчивости выбора исторического окна выполнена повторная перекалибровка модели на сокращённых периодах (1995-2023; 2000-2023; 2005-2023) с повторением РСА-разложения и переоценкой динамики временных коэффициентов (прил. 3, табл. 1). Сопоставимость результатов оценивалась как по параметрам динамики компонент, так и по калибровке прогнозных интервалов в скользящем *backtesting* на горизонте до 5 лет: при переходе от периода 1990-2023 гг. к окнам 2000-2023 и 2005-2023 гг. среднее покрытие 90%-го интервала снижается с 0,68 до 0,48, а его средняя ширина сокращается с 345,0 тыс. человек до 126,4 и 120,9 тыс. человек соответственно (прил. 3, табл. 1). Проверка на укороченных окнах калибровки показывает, что включение ранних лет наблюдений влияет прежде всего на масштаб оценённой неопределённости, то есть на ширину предиктивных интервалов, тогда как медианные прогнозные траектории сохраняют относительную устойчивость. Поэтому период 1990-2023 гг. выбран не только из-за большей длины ряда. Для вероятностного прогноза важно оценивать не одну среднюю траекторию, но и возможный разброс будущих значений. Более длинное окно даёт больше наблюдений для оценки такой неопределённости. Поэтому выбор базового периода дополнительно проверялся через скользящий *backtesting* предиктивных интервалов.

Сравнение окон калибровки показывает, что переход к периоду 2000-2023 гг. действительно меняет структуру данных. По рождаемости возрастает доля

дисперсии, объясняемая первой компонентой, $PC1=98,78\%$. По миграции, наоборот, вклад первой компоненты снижается, а второй возрастает. Это означает, что возрастное распределение миграции хуже описывается одной главной компонентой и требует учёта второй компоненты (прил. 2, табл. 1). Это подтверждает, что выбор исторического окна влияет не только на ширину прогнозных интервалов, но и на саму внутреннюю структуру моделируемых возрастных профилей.

По критериям AIC/BIC модели с $d=1$ иногда дают лучший результат, но это преимущество не является устойчивым. Оно сильнее проявляется при включении 1990-х гг. и ослабевает при их исключении. Поэтому выбор d нельзя считать однозначным для всех компонент модели. В качестве базового окна оставлен период 1990-2023 гг., поскольку для вероятностного прогноза важно сохранить исторический масштаб колебаний. Сокращённые интервалы используются для проверки чувствительности и показывают, как меняются результаты при исключении раннего переходного периода.

После прогнозирования временных коэффициентов выполняется обратный переход от главных компонент к значениям показателя по возрастным группам. Для каждой симуляции d и прогнозного года t расчёт задаётся формулой (3.11):

$$\hat{X}_t^{(d)} = \exp\left(\mu + \sum_{k=1}^K z_{k,t}^{(d)} v_k\right), \quad (3.11)$$

где $\hat{x}_t^{(d)}$ – восстановленный возрастной профиль соответствующего демографического показателя в симуляции d и году t ; μ – средний возрастной профиль; v_k – k -я главная компонента; $z_{k,t}^{(d)}$ – прогнозное значение временного коэффициента k -й компоненты. Экспоненциальное преобразование используется для возврата из логарифмической шкалы в исходную шкалу коэффициентов. Так для каждого показателя формируется множество возможных возрастных траекторий, согласованных с наблюдаемыми историческими закономерностями.

Стохастическая процедура реализована в функции `pca_arima_simulation()`, которая применялась к каждой демографической компоненте. Для смертности и

миграции расчёты выполнены отдельно по полу. Функция выполняет логарифмирование данных, выделение двух главных компонент, оценку авторегрессионных моделей первого порядка AR(1), генерацию случайных отклонений и обратное преобразование в исходную шкалу. Пример реализации приведён в листинге 4:

Листинг 4 – Стохастическое моделирование демографических компонент методом PCA-AR(1)

```
pca = PCA(n_components=min(n_comp, X.shape[1])).fit(X)
scores = pca.transform(X)

for i in range(pca.n_components_):
    model = ARIMA(scores[:, i], order=(1, 0, 0))
    res = model.fit()
    forecast = res.forecast(steps=steps)
    sigma = np.std(res.resid)

    for d in range(draws):
        noise = np.random.normal(0, sigma, size=steps)
        scores_sims[d, :, i] = forecast + noise

def reconstruct(scores):
    return np.exp(np.dot(scores, pca.components_) + pca.mean_)
```

Примечание. Полная версия программного кода представлена в приложении 4.

Поскольку достаточность двухкомпонентного представления была подтверждена выше, дальнейшая процедура строится на моделировании временных коэффициентов первых двух главных компонент. Полученные стохастические реализации рождаемости, смертности и миграции далее используются в Монте-Карло-симуляции численности населения Санкт-Петербурга на горизонте 2024-2045 гг.

3.1.3. Стохастическая когортно-компонентная симуляция (Монте-Карло-модель)

Вероятностный прогноз численности населения Санкт-Петербурга формировался методом Монте-Карло на множестве, $N = 10\,000$ стохастических реализаций. Каждая реализация моделирует развитие половозрастной структуры населения города при действии трёх демографических процессов – рождаемости, смертности и миграции, которые представлены как случайные процессы с исторически наблюдаемой вариабельностью.

В работе использовался годичный шаг прогноза, возрастная сетка представлена пятилетними группами, стартовым населением выступает возрастнополовая структура Санкт-Петербурга на 2023 г., число стохастических реализаций составляет 10 000. Для итоговой интерпретации использовались медиана прогноза, а также 75% и 90% предиктивные интервалы. Для обеспечения воспроизводимости вычислений в программной реализации задано фиксированное начальное значение генератора случайных чисел `pr.random.seed(42)`. Число 42 используется как технический параметр и не влияет на содержательную интерпретацию результатов.

Величина выборки стохастических реализаций N определяется компромиссом между статистической устойчивостью результатов и вычислительной затратностью модели. В практике демографического моделирования количество симуляций, как правило, варьирует от 1 000 до 50 000. Выбранное значение N занимает промежуточное место в этом диапазоне, что обеспечивает сходимость распределений прогнозных показателей по всем возрастным группам [46, 48, 107].

Каждая из 10 000 траекторий начинается с наблюдаемой половозрастной структуры населения Санкт-Петербурга за 2023 год $P_{a,2023}^{(s)}$, где a обозначает возрастную группу (пятилетние когорты от 0 – 4 до 85 +), а $s \in \{F, M\}$ – пол. Это состояние задаёт так называемую точку входа прогноза (jump-off). Дальнейшее обновление структуры населения по годам строится по стандартной когортно-компонентной схеме, но в стохастической формулировке. Для каждой возрастной группы a , пола s и года прогноза t рассчитывается численность следующего года $t + 1$:

$$P_{a,t+1}^{(s)} = A \left[P_{a,t}^{(s)} * S_{a,t}^{(s)} \right] + M_{a,t}^{(s)} + B_{a,t}^{(s)}, \quad (3.12)$$

В этом выражении собраны четыре механизма, которые управляют эволюцией каждой возрастно-половой когорты: дожитие, старение, миграция и рождение новых когорт.

Годовая вероятность дожития $S_{a,t}^{(s)}$ получается из возрастно-половых коэффициентов смертности $ASMR_{a,t}^{(s)}$, сгенерированных ранее методом PCA-AR(1) как стохастические траектории:

$$S_{a,t}^{(s)} = \exp\left(-\frac{m_{a,t}^{(s)}}{1000}\right), \quad (3.13)$$

где $m_{a,t}^{(s)}$ – это смертность в промилле для данного возраста, пола и года.

Такая экспоненциальная запись интерпретирует смертность как интенсивность выбывания в течение года и переводит её в вероятность выживания. Она стандартно используется в стохастических когортно-компонентных моделях и в субнациональных вероятностных прогнозах населения [54].

Для возрастной группы 85 лет и старше прогнозные значения смертности дополнительно корректируются через множитель $OLD_AGE_MORT_MULT = 1,05$: смертность в этой группе увеличивается на 5%, а вероятность дожития немного снижается.

Коррекция для возрастной группы 85+ связана со специальным статусом этой группы в демографических расчётах. В отличие от остальных пятилетних интервалов, группа 85+ является открытой и объединяет возрасты, существенно неоднородные по риску смертности. При использовании сглаженных временных рядов это создаёт систематическое смещение в сторону завышения выживаемости на дальнем горизонте прогноза. Поэтому поправка +5% выступает не как произвольное ужесточение сценария, а как консервативная bias-correction для открытого возрастного интервала, позволяющая избежать накопления оптимистической ошибки в старших когортных возрастах.

Величина поправки выбрана на основе анализа типичных ошибок долгосрочного моделирования смертности в старших возрастах. Как отмечает Н. Кейльман (Niko Keilman), при экстраполяции сглаженных временных рядов модели часто переоценивают выживаемость пожилых когорт на горизонтах свыше 15 лет [91]. Это связано с тем, что статистические колебания смертности в

возрастах 80+ выражены сильнее, а наблюдаемые уровни подвержены случайным флуктуациям из-за малых численностей.

В целях компенсации систематического смещения вводилась умеренная корректировка. Принятое в данном исследовании значение +5 % (коэффициент $OLD_AGE_MORT_MULT = 1,05$) соответствует среднему уровню поправки.

Это подтверждено результатами региональных норвежских и австралийских прогнозах, где городские и окружные оценки смертности в старших возрастах оказались статистически нестабильными [91; 109]). В дополнение для каждой возрастной группы в настоящей работе введён верхний предел допустимых значений смертности (массив CAP_BY_AGE): не выше 35‰ для возрастов младше 50 лет, не выше 55‰ для 50-64 лет, 95‰ для 65-74 лет, 160‰ для 75-79 лет, 260‰ для 80-84 лет и 340‰ для 85+. Эти ограничения не «улучшают» прогноз визуально, а отсекают единичные выбросы, которые появляются, когда стохастическая компонента смертности умножается на малочисленную когорту и создаёт нереалистичный скачок.

Рассмотрим старение когорты. Матричный оператор A использовался для перехода населения между пятилетними возрастными группами при годовичном шаге прогноза. Сначала для каждой возрастно-половой группы применялась соответствующая вероятность дожития, поэтому различия в смертности учитывались на предыдущем шаге расчёта. После этого оставшаяся численность распределялась между возрастными группами: 4/5 когорты оставались в прежней пятилетней группе, а 1/5 переходила в следующую. Такое соотношение не означает одинаковую вероятность смерти или дожития для всех возрастов, а используется как годовой способ передвижения населения внутри пятилетней возрастной сетки. Для группы 85 лет и старше дальнейшее возрастное движение не детализировалось: все лица, достигшие этого возраста, оставались в группе 85+, так как следующая возрастная группа в используемой сетке не предусмотрена.

Что касается миграции, то в модели миграция вводится не как одна фиксированная «экспертная поправка», а как стохастический процесс с возрастно-половой структурой. Это ключевое отличие субнациональных вероятностных

прогнозов от классических вариантов высокого/низкого типа. Вначале методом PCA-AR(1) были получены стохастические годовые профили чистой миграции для женщин и мужчин по возрасту, то есть величины $M_{a,t}^{(F)}$ и $M_{a,t}^{(M)}$ до корректировок.

Далее в модели миграции применялись два уровня уточнения. Первый уровень – эмпирическое формирование возрастно-полового профиля миграции. Для этого рассчитывалось распределение прибывших по возрасту отдельно для женщин и мужчин за окно 2014-2023 гг. Окно 2014-2023 использовалось здесь не для оценки общего уровня миграции и не для задания её долгосрочной динамики, а только для стабилизации возрастно-половой формы базового миграционного профиля. Сами уровни и стохастическая динамика миграции оценивались на полном историческом ряде 1990-2023 гг. в блоке PCA-AR(1).

Второе уточнение состоит в том, что эти распределения усредняются с помощью экспоненциально-взвешенного сглаживания с параметром $\lambda = 0,8$, то есть последние годы получают больший вес, чем начало окна [86]. Такая экспоненциальная взвешенность используется в субнациональных прогнозах, когда целью является уловить «текущую структуру миграции», а не среднюю структуру за десятилетия [102, 111].

В работе сравнивались несколько способов получения возрастных весов миграции, включая скользящее среднее и среднее по полному периоду. Качество сопоставлялось по среднеквадратической ошибке и средней абсолютной процентной ошибке $MAPE$. Сравнение показало, что окно 2014-2023 гг. с параметром $\lambda = 0,8$ лучше всего воспроизводит фактическое распределение прибывших по возрасту к 2023 г. Отклонение по возрастным группам не превышает 3%. Это означает, что веса $w_F(a)$ и $w_M(a)$ отражают сложившееся распределение миграционного притока в Санкт-Петербург, где основной вклад приходится на трудоспособные возраста, но сохраняется и небольшой вклад старших возрастов, около 4-6% прибывших в возрасте 70 лет и старше.

Далее эти веса использовались для того, чтобы прогнозируемый миграционный поток сохранял реалистичный возрастно-половой состав. В модели он задаётся формулой (3.14):

$$M_{a,t}^{(s)} = \text{clip} \left[r_{a,t}^{(s)} P_{a,t}^{(s)} w_s(a), -C, C \right], \quad (3.14)$$

где $r_{a,t}^{(s)}$ – стохастически сгенерированная интенсивность миграции, полученная из PCA-AR(1)), $P_{a,t}^{(s)}$ – текущая численность соответствующей когорты, $w_s(a)$ – эмпирические возрастные веса для пола s , C – верхнее ограничение миграционного сальдо на возрастную группу. В модели было принято $C = 8\,000$ человек в год на одну пятилетнюю возрастную-половую группу. Это значение не отражает фактический ежегодный приток, а выступает техническим верхним ограничением для миграционного прироста. Такой «кап» предотвращает появление заведомо нереалистичных всплесков миграции в одном узком возрасте (например, условный «приток» сразу 25 тысяч мужчин 25-29 лет за один год), которые могут возникнуть в стохастической модели, но не наблюдаются даже в крупнейших агломерациях. Ограничение пиковых миграционных притоков с помощью «потолка» для возрастную-половую группы рекомендуется именно в субнациональных прогнозах, потому что на уровне города/округа единичные всплески могут быть статистическим шумом, а не устойчивым трендом.

Второе ограничение на уровеньный «миграционный bump»: вводится калибровочный сдвиг в трудоспособных возрастах. В модели задаётся дополнительный чистый миграционный приток $MIG_NET_BUMP = 8\,000$ человек в год, распределённый по возрастным группам 20-44 года с пиком в возрастном интервале 25-34 лет. Распределение формируется треугольной весовой функцией, то есть приток максимален в ядре модального трудового возраста и линейно снижается к его краям. Этот приток делится по полу с коэффициентами 0,48 для женщин и 0,52 для мужчин, что отражает эмпирически наблюдаемую некоторую «мужскую избыточность» трудовой миграции. По сути, данный элемент играет роль «реанкеринга» долгосрочного уровня миграции: вместо того чтобы полностью доверять последнему наблюдению, модель удерживает приток квалифицированного трудоспособного населения на уровне, согласованном с реалистичными среднесрочными ожиданиями городской экономики. Аналогичная

логика вводится в прикладных прогнозах городов и округов [110, 117], где миграция признаётся основным источником расхождений прогноза и факта.

Относительно рождаемости: число рождений в году t определяется на основе возрастных коэффициентов рождаемости женщин $ASFR_{a,t}$ (в промилле) и численности женского населения в соответствующих репродуктивных когортах:

$$B_t = \sum_{a \in A_f} \frac{ASFR_{a,t} * P_{a,t}^{(F)}}{1000}, \quad (3.15)$$

где $A_f = \{15-19, 20-24, \dots, 45-49\}$.

Все новорождённые зачисляются в самую младшую возрастную группу (0-4 года). Далее новорождённые разбиваются по полу с использованием соотношения полов при рождении $SRB = 1,06$, что соответствует 106 мальчиков на 100 девочек. Значение $SRB = 1,06$ взято по данным территориальной статистики (Петростат) за последние годы и согласуется с типичным уровнем для России. Это означает:

$$p_{male} = \frac{SRB}{1 + SRB}, \quad (3.16)$$

$$p_{female} = \frac{1}{1 + SRB}, \quad (3.17)$$

соответственно,

$$P_{0,t+1}^{(M)} = B_t \cdot p_{male}, \quad (3.18)$$

$$P_{0,t+1}^{(F)} = B_t \cdot p_{female}, \quad (3.19)$$

Фиксация SRB как стабильного параметра на горизонте в 20 лет принадлежит к стандартным допущениям вероятностных когортно-компонентных прогнозов: модель допускает, что рождаемость колеблется по возрастным коэффициентам, но не переоценивает вероятность резкого сдвига соотношения полов при рождении за короткий интервал.

Все описанные шаги были выполнены последовательно для каждого года прогноза (2024, 2025, ..., 2045) и для каждой симуляции $d = 1, \dots, N$. Код, реализующий этот цикл, представлен в листинге 5. В нём на каждом шаге выполняются: (1) применение вероятностей дожития и матрицы старения A ; (2) добавление миграции с эмпирически заданной возрастно-половой формой и

ограничением по максимуму; (3) добавление уровневого миграционного «bump» в трудоспособном возрасте; (4) генерация рождений и разнесение новорождённых по полу с использованием $SRB = 1,06$.

Листинг 5 – Фрагмент реализации Монте-Карло-симуляции когортно-компонентного прогноза

```
for d in range(N):
    popF = base_F.copy()
    popM = base_M.copy()

    for t in range(T):
        mortF = mort_models["F"]["sims"][d, t, :]
        mortM = mort_models["M"]["sims"][d, t, :]
        migF = mig_models["F"]["sims"][d, t, :]
        migM = mig_models["M"]["sims"][d, t, :]
        asfr = fert_model["sims"][d, t, :]

        survF = surv_prob_exp_by_age(mortF)
        survM = surv_prob_exp_by_age(mortM)

        popF = age_population(popF, survF)
        popM = age_population(popM, survM)

        popF += shape_migration_sex(migF, "F", cap=MIG_GROUP_CAP)
        popM += shape_migration_sex(migM, "M", cap=MIG_GROUP_CAP)

        births = births_from_asfr(popF, asfr)
        popF[0] += births * p_fem
        popM[0] += births * p_male
```

Примечание. Полная версия программного кода представлена в приложении 4.

Распределение числа рождений по полу входит в обновлённые возрастно-половые матрицы населения, которые сохраняются для каждой симуляции и каждого года прогноза.

После завершения моделирования по всем $N = 10\,000$ стохастическим реализациям формируется распределение будущей численности населения Санкт-Петербурга для каждого года прогноза. Вычисляются медианное значение $\tilde{P}_t$, а также нижняя и верхняя границы предиктивного интервала:

$$P_t^{(10)} = q_{0.10}(P_{d,t}), \quad (3.20)$$

$$P_t^{(90)} = q_{0.90}(P_{d,t}), \quad (3.21)$$

где $P_{d,t} = \sum_{a,s} P_{a,t}^{(s,d)}$ – суммарная численность населения по симуляции d .

Тем самым прогноз даётся не как одна линия, а как ансамбль предиктивных интервалов. Статистическая интерпретация медианы – как наиболее типичный исход при допущениях модели, граница 90 % – диапазон значений, совместимых с исторической волатильностью рождаемости, смертности и миграции в мегаполисе.

Все ключевые элементы вероятностной схемы – выбор двух главных компонент, использование парсимоничной динамики AR(1), число стохастических реализаций, способ задания сценарных шоков и корректировки для старших возрастов – определялись не из соображений вычислительного удобства как такового, а исходя из критерия устойчивости прогноза. В рамках исследования предпочтение отдавалось тем решениям, которые сохраняют интерпретируемость демографических компонент, не переобучаются на частные колебания исторического периода и обеспечивают воспроизводимую оценку не только медианной траектории, но и предиктивной неопределённости (прил. 4). Именно в этом смысле предложенная модель представляет собой не набор технических процедур, а согласованную модифицированную автором методику вероятностного когортно-компонентного прогнозирования мегаполиса.

3.1.4. Калибровка предиктивных интервалов по результатам ретроспективного моделирования (backtest)

Ключевым элементом реализации является калибровка доверительных интервалов на основе ретроспективного моделирования (backtesting) [81, 110]. Исторический период разбивается на последовательность базовых годов; для каждого такого года строится вероятностный прогноз на заданный горизонт, после чего проверяется, как часто фактическая численность населения попадает в требуемые 75%-ный или 90%-ный доверительные интервалы.

Одной из главных характеристик валидности прогноза является показатель эмпирического покрытия $C_p(h)$ для горизонта h . Этот показатель рассчитывается как доля попадания фактических значений в границы доверительного интервала с номинальным уровнем p :

$$C_p(h) = \frac{1}{|O|} \sum_{o \in O} I(P_{o+h}^{fact} \in [P_{o+h}^{L,p}, P_{o+h}^{U,p}]), \quad (3.22)$$

где $|O|$ – число стартовых лет; $I\{\cdot\}$ – индикатор попадания; $P^{L,p}$ и $P^{U,p}$ – нижняя и верхняя границы интервала.

Если эмпирическое покрытие $C_p(h)$ существенно ниже целевого уровня, интервальные оценки расширяются путём умножения половины доверительного интервала на калибровочный множитель:

$$k_p(h) = \frac{p}{C_p(h)}, p \in \{0,75; 0,90\}, \quad (3.23)$$

Обычно этот множитель ограничен диапазоном значений $[1,0; 3,5]$. Например, для горизонта $h = 5$ лет при $C_{0,90}(5) = 0.72$ получаем $k_{0,90}(5) \approx 1.25$, что указывает на необходимость увеличить полуинтервал примерно на 25%.

После применения таких множителей по всем горизонтам достигается эмпирическое покрытие, близкое к 0,90 для 90%-ных и к 0,75 для 75%-ных интервалов. Как показывает опыт региональных прогнозов численности населения, получаемые таким образом характеристики сопоставимы с результатами backtest-калибровки для региональных прогнозов в Новой Зеландии и в округах США [81, 117].

3.1.5. Построение итоговых предиктивных интервалов и визуализация результатов

После выполнения процедуры backtest-калибровки стохастическая модель получила набор корректирующих множителей $k_H^{(p)}$, которые отражают зависимость между длиной прогноза и требуемой шириной предиктивных интервалов. На этом этапе были сформированы окончательные вероятностные границы прогноза и выполнена их графическая интерпретация.

Основой послужила матрица стохастических реализаций численности населения, полученная в результате Монте-Карло симуляций из раздела 3.1.3, где каждая траектория $P_t^{(d)}$ описывает возможную динамику населения в год t по результатам моделирования компонент рождаемости, смертности и миграции. Для каждого прогнозного года были вычислены медианные значения и процентильные границы исходного распределения:

$$\begin{aligned} Med_t &= median(P_t^d), & P_t^5 &= Q_{0.05}(P_t^d), \\ P_t^{95} &= Q_{0.95}(P_t^d), & P_t^{12.5} &= Q_{0.125}(P_t^d), & P_t^{87.5} &= Q_{0.875}(P_t^d), \end{aligned} \quad (3.24)$$

где Q_q – q -й квантиль эмпирического распределения. Эти границы формируют «сырые» интервалы до применения калибровочных поправок.

Чтобы обеспечить согласованность 90%-го и 75%-го предиктивных интервалов с ретроспективно оценёнными уровнями неопределённости на каждом горизонте $H = t - t_0 + 1$ (где $t_0 = 2024$ – первый прогнозный год) интервалы масштабируются с использованием множителей $k_{90}(H)$ и $k_{75}(H)$, рассчитанных на основе эмпирического покрытия. Калиброванные границы интервалов рассчитывались путём масштабирования полуинтервала от центра соответствующего исходного интервала:

$$m_t^{90} = \frac{P_t^{95} + P_t^5}{2}, \quad h_t^{90} = \frac{P_t^{95} - P_t^5}{2} \quad (3.25)$$

$$\tilde{P}_t^5 = m_t^{90} - k_{90(H)} \cdot h_t^{90}, \quad \tilde{P}_t^{95} = m_t^{90} + k_{90(H)} \cdot h_t^{90} \quad (3.26)$$

$$m_t^{75} = \frac{P_t^{87.5} + P_t^{12.5}}{2}, \quad h_t^{75} = \frac{P_t^{87.5} - P_t^{12.5}}{2} \quad (3.27)$$

$$\tilde{P}_t^{12.5} = m_t^{75} - k_{75(H)} \cdot h_t^{75}, \quad \tilde{P}_t^{87.5} = m_t^{75} + k_{75(H)} \cdot h_t^{75} \quad (3.28)$$

где m_t^{90} и m_t^{75} – центры исходных 90- и 75-процентных интервалов; h_t^{90} и h_t^{75} – соответствующие полуинтервалы; $k_{90}(H)$ и $k_{75}(H)$ – калибровочные множители для горизонта прогноза H . Величины $\tilde{P}_t^5$ и $\tilde{P}_t^{95}$ обозначают нижнюю и верхнюю границы 90%-го предиктивного интервала, а $\tilde{P}_t^{12.5}$ и $\tilde{P}_t^{87.5}$ – границы 75%-го предиктивного интервала.

В программной реализации (листинг 6) этот шаг записывается так:

Листинг 6 – Формирование медианной траектории и предиктивных интервалов прогноза

```
# Формирование медианной траектории и интервальных границ прогноза

p5 = np.percentile(totAll, 5, axis=0)
p95 = np.percentile(totAll, 95, axis=0)
p12 = np.percentile(totAll, 12.5, axis=0)
p87 = np.percentile(totAll, 87.5, axis=0)
median_pop = np.median(totAll, axis=0)
```

```
total_df = pd.DataFrame({
    "Year": YEARS,
    "Median": median_pop,
    "P5": p5,
    "P95": p95,
    "P12_5": p12,
    "P87_5": p87
})

# Применение калибровки предиктивных интервалов по горизонту прогноза

total_df_CAL = apply_calibration_by_horizon(
    total_df,
    k_byH["90"],
    k_byH["75"],
    years=YEARS
)
```

Примечание. Полная версия программного кода представлена в приложении 4.

Полученная итоговая расчётная таблица содержит медианные значения и калиброванные предиктивные интервалы численности населения Санкт-Петербурга для каждого года прогнозного периода 2024-2045 гг. Интервалы уточнены с использованием калибровочных множителей $k_{90}(H)$ и $k_{75}(H)$, рассчитанных по результатам ретроспективной проверки.

Полные годовые значения медианной траектории и калиброванных предиктивных интервалов представлены в приложении 7 (табл. 1).

3.2. Результаты базового вероятностного прогноза численности и возрастно-половой структуры населения Санкт-Петербурга

История Санкт-Петербурга подтверждает необходимость особого подхода к демографическому прогнозированию. Город неоднократно переживал демографические катастрофы, граничащие с «клинической смертью»: резкое сокращение численности в 1919-1920 годах, тяжелейшие потери в период блокады 1941-1943 годов, оживление в 1970-х гг. на фоне сохраняющихся ограничений регистрации, затяжной спад начала 1990-х, когда миграционное сальдо стало отрицательным и в 1993 году достигло минус 16 тысяч человек.

В 2000-е годы, Санкт-Петербург достиг состояния относительной демографической стабилизации с темпами естественного прироста населения близкими к нулю, миграционными притоками, которые не всегда компенсируют

естественные потери. От отправной точки – базового 2023 года – приводятся результаты стохастического прогноза, описывающий вероятные траектории численности и возрастно-половой структуры населения города до 2045 года.

В соответствии с медианной траекторией численность населения Санкт-Петербурга в середине прогнозного горизонта достигает 5,78 млн человек, а к 2045 году плавно снижается до 5,62 млн человек.

Диапазон неопределённости остаётся умеренным: 90-процентный предиктивный интервал охватывает значения от 5,22 до 6,04 млн человек, то есть размах около 0,8 млн. Такой уровень разброса является ожидаемым для долгосрочных стохастических прогнозов мегаполисов и отражает совокупное влияние вариаций рождаемости, смертности и миграции (рис. 16).

Поведение кривой указывает на переход Санкт-Петербурга от фазы стремительного восстановления демографических процессов до уровня советского периода к состоянию демографического равновесия, когда естественные компоненты прироста постепенно теряют значение, а миграция становится основным компенсирующим фактором.

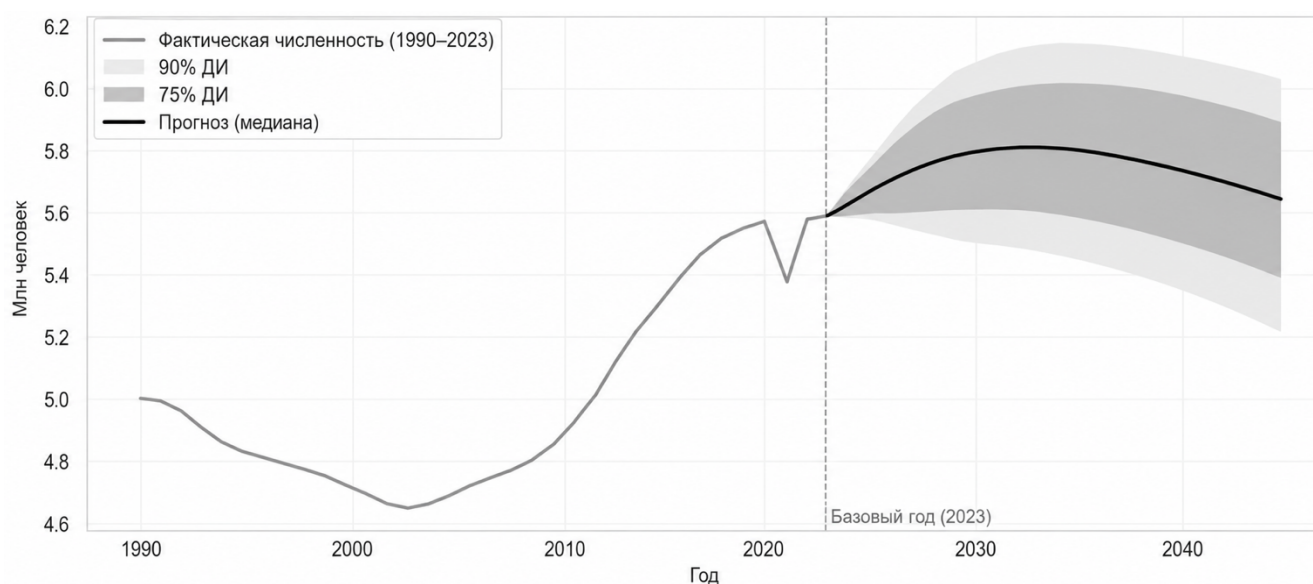


Рисунок 16 – Фактическая и прогнозная численность населения Санкт-Петербурга, 1990-2045 гг.

Первая половина прогнозного периода несёт в себе эхо демографического подъёма, начавшегося в 2007 г. и достигшего максимума по числу родившихся в 2016 г. Поколения, родившиеся в эти годы, к 2030-м гг. постепенно входят в

активный репродуктивный возраст [59]. С учетом того, что средний возраст матери при рождении первого ребёнка в Петербурге составляет 27,6 лет [59], именно эти когорты формируют кратковременное увеличение числа рождений, своеобразный «второй гребень» демографической волны в 2030-2036 гг.

Если сопоставить результаты стохастической модели с официальными оценками Росстата, видно, что направления динамики совпадают, но уровни несколько различаются (рис. 17).

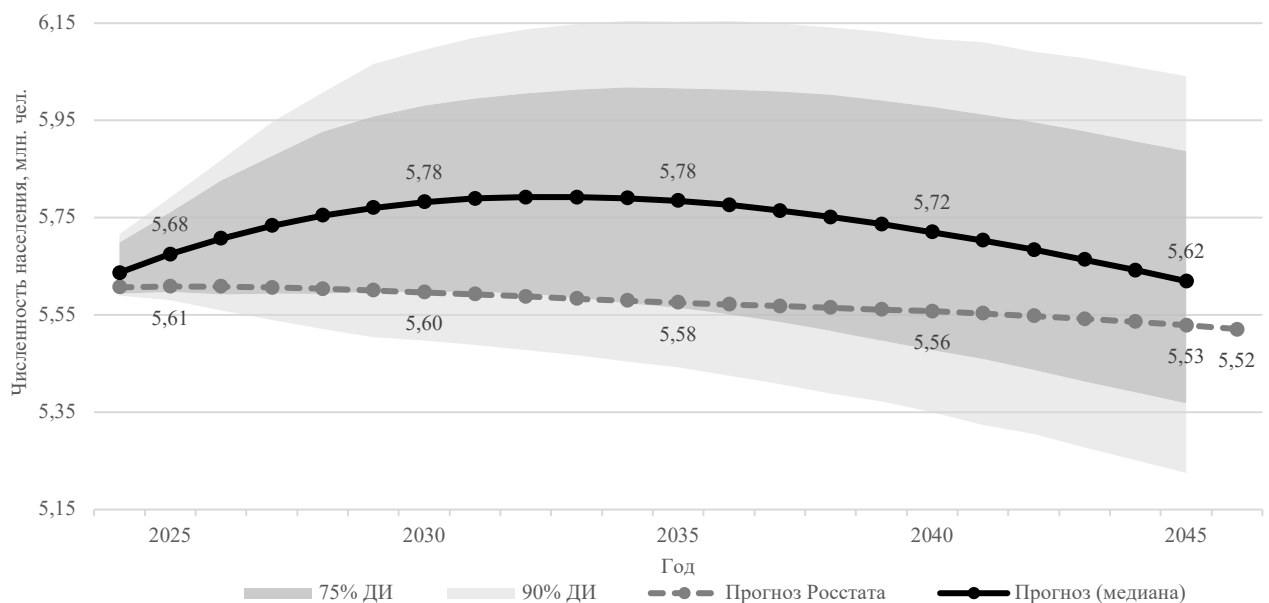


Рисунок 17 – Сравнение стохастического прогноза и прогноза Росстата для Санкт-Петербурга, 2024-2045 гг.

К 2045 году по медианной траектории стохастического прогноза численность населения Санкт-Петербурга оценивается на уровне 5,62 млн человек, тогда как инерционный сценарий Росстата дает оценку 5,52 млн человек, т.е. на 100 тыс. человек меньше. Разрыв между оценками сравнительно невелик, что подтверждает сопоставимость результатов. Данное расхождение обусловлено расчётно-методическими особенностями.

Во-первых, исходная база численности в настоящем исследовании использована по полу и возрасту на основе среднегодовой численности населения за 2023 год, тогда как в расчётах Росстата за базу принимается численность на начало года (01.01.2023). Даже такая, казалось бы, незначительная разница в

моменте фиксации – порядка нескольких десятков тысяч человек – приводит к расхождению в стартовой точке, которое на длинном прогнозном горизонте аккумулируется в заметную разницу итоговых оценок.

Во-вторых, миграционный компонент стохастической модели реализован более детально и статистически обоснованно – по полу, возрасту и году, с применением экспоненциально-взвешенного сглаживания последних наблюдений ($\lambda = 0,8$) и последующего стохастического моделирования интенсивностей. Такой подход позволяет уловить актуальную возрастную структуру миграции, а не усреднённые многолетние профили. По результатам валидации через показатели среднеквадратичной и средней абсолютной процентной ошибки (MAPE) установлено, что ошибка воспроизведения возрастного рисунка миграции не превышает 2-3 %, что подтверждает высокую точность аппроксимации. Такая постановка согласуется с современными подходами к вероятностному прогнозированию миграции, где миграционные потоки моделируются не как фиксированное сальдо, а как вероятностные траектории, дифференцированные по полу и возрасту [74, 84].

Кроме того, в модели реализован уровневой «bump»-эффект – дополнительный чистый приток трудоспособного населения, распределённый по возрастам 20-44 года с пиком в интервале 25-34 лет. Этот элемент отражает калибровочный сдвиг, обусловленный устойчивым положительным миграционным балансом Петербурга как ведущего центра занятости и образования [37, 53]. На протяжении прогнозного горизонта bump-эффект обеспечивает сохранение умеренного притока экономически активных когорт, тогда как в сценариях Росстата этот поток постепенно затухает, предполагая ослабление миграционного притяжения мегаполиса.

Средний вариант прогноза Росстата на всём прогнозном периоде попадает в предиктивный интервал авторской модели. К концу периода траектории сближаются, что отражает близость общего направления расчётов. В 2025 г. медианная оценка авторского прогноза составила 5,68 млн человек, что на 1,2% выше среднего варианта Росстата, равного 5,61 млн человек. В середине

прогнозного периода, в 2030-2035 гг., различие становится заметнее. Авторская медианная траектория держится около 5,78 млн человек, тогда как прогноз Росстата находится на уровне 5,57-5,60 млн человек. К 2045 г. разрыв сокращается, при этом авторская модель даёт 5,62 млн человек, а официальный прогноз составляет 5,52 млн человек.

Росстат исходит из более сдержанной оценки будущего миграционного притока. В стохастическом прогнозе миграция сохраняет положительный вклад, хотя и без резкого роста. Это показывает, что расхождение между расчётами связано прежде всего с разной оценкой способности Петербурга удерживать миграционное притяжение.

При этом основной сдвиг происходит не столько в общей численности, сколько в возрастно-половой структуре населения. Прогноз показывает постепенное старение Санкт-Петербурга, рост старших возрастных групп и сравнительно слабое пополнение младших возрастов (рис. 18).

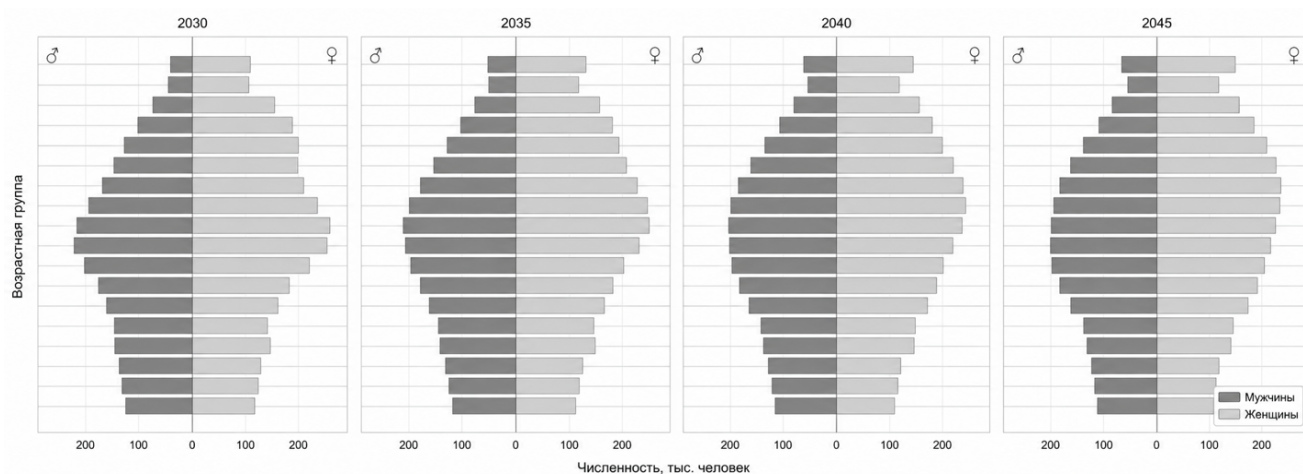


Рисунок 18 – Прогноз возрастно-половой структуры населения Санкт-Петербурга:
2030 – 2045 гг.

Возрастно-половые пирамиды (рис. 18) показывают постепенное старение населения Санкт-Петербурга. Младшие возраста сужаются, старшие группы расширяются, а численность трудоспособных возрастов пока остаётся относительно высокой. До 2030 г. сохраняется влияние демографической волны начала 2000-х годов, поэтому наиболее многочисленными остаются группы 30-39 и 40-44 лет.

К 2035 г. возрастно-половая структура смещается к старшим возрастам. Численность детей и подростков снижается, а группа 50-64 лет увеличивается за счёт многочисленных поколений 1980-х годов. Доля населения 65 лет и старше становится близкой к уровню крупных европейских городов с выраженным старением. Пирамида приобретает более вытянутую форму, а различия между младшими и старшими возрастными группами постепенно сглаживаются. К 2045 г. доля жителей 65 лет и старше достигает 22-23%, а доля детей 0-14 лет снижается до 12%. Население 15-64 лет остаётся наиболее многочисленной частью городской структуры, однако его доля сокращается. Для Санкт-Петербурга это означает переход к более зрелой возрастной структуре, где устойчивость зависит от баланса между трудоспособными, младшими и старшими поколениями.

Половые различия постепенно сглаживаются. В старших возрастах женщины по-прежнему преобладают, но разрыв между полами сокращается из-за улучшения выживаемости мужчин. В возрастах до 60 лет соотношение мужчин и женщин близко к паритетному, тогда как после 65 лет сохраняется выраженное преобладание женщин: 1744 женщины на 1000 мужчин (табл. 13).

Таблица 13 – Возрастная структура населения Санкт-Петербурга, прогноз до 2045 г.

Год	0-14 лет	15-64 года	65 лет и старше	Всего, тыс. чел.	0-14, %	15-64, %	65+, %
2030	783	3 852	1 178	5 814	13,5	66,3	20,3
2035	752	3 876	1 232	5 860	12,8	66,1	21
2045	699	3 788	1 293	5 780	12,1	65,5	22,4

Из данных видно, что доля населения трудоспособного возраста снижается умеренно, с уровня около двух третей в 2030 г. до 65% к 2045 г. Одновременно увеличивается доля старших возрастов. Особенно заметно растут группы 70 лет и старше, что связано с постепенным снижением смертности и увеличением продолжительности жизни. Медианные значения прогнозной возрастно-половой структуры населения Санкт-Петербурга по реперным годам представлены в приложении 8 (табл. 1-2).

При этом старение населения не означает немедленного исчерпания трудового потенциала города. У Санкт-Петербурга сохраняется демографическое

окно возможностей, поскольку многочисленные поколения, родившиеся в 1980-2000-е годы, ещё остаются в активных возрастах. Именно их участие в экономике будет во многом определять, насколько город сможет адаптироваться к росту нагрузки со стороны старших возрастов.

В ближайшие два десятилетия для Санкт-Петербурга важен не только сам прирост населения, но и то, как используется имеющийся человеческий потенциал. Значение будут иметь продление активной жизни, участие женщин и старших возрастов в занятости, а также развитие социальной среды, позволяющей дольше сохранять экономическую и общественную активность. Поддержание этого баланса между демографической зрелостью и экономической активностью становится ключевым условием устойчивого развития Петербурга в долгосрочной перспективе.

3.3. Перспективные траектории компонент воспроизводства населения

Расхождение итоговых траекторий численности определяется различием гипотез, заложенных в динамику компонент воспроизводства. Стартовый шаг после базового года не интерпретируется как самостоятельная содержательная гипотеза⁴⁴. По итоговой численности населения средний вариант Росстата на всём прогнозном горизонте остаётся внутри доверительного коридора вероятностной модели.

По СКР⁴⁵ к 2030 г. различие минимально: медиана авторского прогноза составляет 1,325 против 1,334 в среднем варианте Росстата. Далее контуры расходятся: Росстат задаёт повышение до 1,493 к 2040 г. и 1,517 к 2045 г., тогда как в вероятностной модели медиана снижается до 1,239 и 1,218. Здесь проявляется различие самой прогнозной конструкции: PCA-AR(1) воспроизводит инерционную

⁴⁴ В стационарной авторегрессионной модели первого порядка AR(1) одношаговый прогноз после базового года в значительной мере определяется оценёнными безусловным средним процесса и коэффициентом инерции. При $|\varphi| < 1$ возникает механический эффект *mean reversion*: условное ожидание следующего значения $\mathbb{E}(x_{t+1} | x_t) = \mu + \varphi(x_t - \mu)$, поэтому прогнозируемое значение x_{t+1} смещается от последнего наблюдения x_t в сторону безусловного среднего μ . Если ряд неоднороден и последние наблюдения нестабильны, первый прогнозный шаг может отражать прежде всего свойства статистической спецификации, а не содержательную демографическую гипотезу. Поэтому интерпретация траекторий переносится на средне- и долгосрочные реперные горизонты.

⁴⁵ СКР – суммарный коэффициент рождаемости. Он показывает сколько в среднем родила бы одна женщина на протяжении всего репродуктивного периода (т.е. от 15 до 50 лет) при сохранении в каждом возрасте уровня рождаемости того года, для которого вычисляется показатель.

динамику возрастных коэффициентов рождаемости, не вводя нормативно заданного восходящего тренда. Перестройка возрастного профиля рождений носит преимущественно структурный характер: возрастает вклад групп 20-24 и 25-29 лет, тогда как вклад возрастов 35-39 лет снижается. Это отражает прежде всего изменение численности женских когорт в репродуктивных возрастах, а не «возврат» к раннему материнству⁴⁶. В абсолютном выражении медианное число рождений сокращается с 46,0 тыс. в 2030 г. до 42,5 тыс. в 2040 г. и 41,4 тыс. к 2045 г.

По миграционной компоненте Росстат исходит из последовательного снижения миграционного прироста: с 18,1 тыс. человек в 2030 г. до 16,2 тыс. в 2040 г. и 15,1 тыс. в 2045 г. В вероятностной модели медианный чистый миграционный прирост также уменьшается, но остаётся заметно выше – 34,0 тыс., 25,2 тыс. и 23,4 тыс. человек соответственно. В обоих подходах к концу горизонта миграция ослабевает как источник роста, однако в стохастическом варианте её компенсирующая роль по отношению к естественной убыли сохраняется дольше; различие уровней связано и с иной калибровкой базовой траектории, и с различием способа задания миграционных предпосылок⁴⁷.

В смертности контраст проявляется как в траекториях ОПЖ⁴⁸, так и в числе умерших. В авторском прогнозе к 2035 г. медиана ОПЖ составляет 79,17 года у женщин и 71,19 года у мужчин, к 2045 г. – 78,60 и 70,68 года; то есть начальный рост не переходит в устойчивую траекторию улучшения. В среднем варианте Росстата заложено более выраженное повышение ОПЖ: 82,89 и 76,63 года к 2035 г. и 84,54 и 79,22 года к 2045 г. В тех же единицах различие фиксируется и по числу

⁴⁶ Под многочисленными когортами понимаются поколения, сформированные на волне подъёма рождаемости, начавшегося с 2007 г. на фоне введения мер демографической поддержки (в частности, материнского капитала). В Санкт-Петербурге число родившихся увеличилось с 43,3 тыс. в 2007 г. до 55,6 тыс. в 2010 г. и 72,7 тыс. в 2016 г. По мере того, как эти поколения входят в активный репродуктивный возраст на прогнозном горизонте, их численность структурно усиливает вклад возрастных групп 20-24 и 25-29 лет в общий объём рождений.

⁴⁷ В методологии Росстата сценарные переменные задаются на уровне агрегатов (для рождаемости – СКР и средний возраст матери; для смертности – ОПЖ и коэффициент младенческой смертности; для миграции – абсолютные числа прибывших и выбывших), затем преобразуются в возрастно-половые коэффициенты и включаются в расчёт по процедуре «передвижки по возрастам»; сценарии уточняются раз в два года.

⁴⁸ ОПЖ – Ожидаемая продолжительность жизни. Показатель характеризует число лет, которое в среднем предстояло бы прожить одному человеку из некоторого гипотетического поколения при условии, что на протяжении всей жизни этого поколения уровень смертности в каждом возрасте останется таким, как в год, для которого вычислен показатель.

умерших: по Росстату 65,9 тыс. человек в 2030 г., 70,2 тыс. в 2040 г. и 73,2 тыс. в 2045 г.; по вероятностной модели – 79,7 тыс., 94,0 тыс. и 96,8 тыс. человек. Медианные значения прогнозных возрастных коэффициентов рождаемости, смертности и миграционной компоненты, а также годовые значения родившихся, умерших и чистой миграции представлены в приложении 9 (табл. 1-4).

Различие носит системный характер: в PCA-AR(1) снижение возрастных коэффициентов смертности оказывается умеренным и не компенсирует ускоряющееся старение возрастной структуры, тогда как официальный средний вариант предполагает более выраженное снижение смертности, частично сглаживающее структурный эффект. Подобно рождаемости, динамика смертности здесь определяется прежде всего возрастной структурой: укрупнение старших когорт усиливает «вес» возрастов с высокой смертностью и тем самым поднимает общий показатель.

3.4. Сценарный анализ в вероятностном прогнозе населения Санкт-Петербурга

В дополнение к базовому стохастическому прогнозу была реализована система сценарных расчётов, позволяющая оценить диапазон возможных траекторий численности населения Санкт-Петербурга в зависимости от внешних социально-экономических условий и возможностей города. К ним относятся состояние рынка труда, уровень заработной платы и доступность жилья, определяющие миграционную привлекательность Санкт-Петербурга [53]. При интерпретации сценариев рождаемости важно учитывать, что влияние демографической политики не является универсальным и зависит от институционального и ценностного контекста. Резервом повышения эффективности мер демографической политики является их «своевременная адаптация и локальная привязка к меняющимся и разнообразным потребностям семей с детьми» [18]. Сценарный подход дополняет вероятностную когортно-компонентную модель, отражая не только статистическую вариацию компонент, но и структурные сдвиги, обусловленные изменениями демографической политики, миграционных потоков и социально-экономических факторов.

Сценарные допущения позволяют моделировать последствия возможных изменений траектории демографического развития, выходящих за пределы статистической инерции. Такие возможности могут открыться перед Санкт-Петербургом в случае улучшения транспортных коммуникаций (более активного введения новых станций метро, увязки метрополитена с ж/д сообщением), создания существенного числа высокотехнологичных рабочих мест в области нанотехнологий, фармакологии и прочих, реализации экологических мероприятий по обеспечению комфортной среды обитания, развития инфраструктуры (строительства спортивных сооружений различного назначения, плавательных бассейнов и т.д.).

Сценарии С1-С4 охватывают конфигурации демографического развития – от благоприятного до неблагоприятного. Сценарный анализ основан на корректировке возрастных коэффициентов рождаемости, возрастных коэффициентов смертности мужчин и женщин, а также дополнительного годового миграционного прироста, распределяемого по возрастно-половым группам. Выбор четырёх сценариев связан с тем, что базовая траектория уже представлена в вероятностном прогнозе, а сценарный блок предназначен для оценки отклонений от неё. Выбор четырёх сценариев связан с логикой сценарного анализа, предполагающей рассмотрение нескольких альтернативных вариантов будущего развития при разных сочетаниях исходных предположений [104].

В сценарии С1 задаются повышение возрастных коэффициентов рождаемости на 5%⁴⁹, ежегодное снижение возрастных коэффициентов смертности мужчин на 1,2 % и женщин на 0,9 %, а также дополнительное годовое чистое сальдо миграции 35 тыс. человек. В сценарии С2 соответствующие параметры составляют +3 %, –0,8 %, –0,6 % и 25 тыс. человек. В сценарии С3 задаются снижение возрастных коэффициентов рождаемости на 2 %, ежегодное снижение смертности мужчин на 0,3 % и женщин на 0,2 %, а также дополнительное годовое чистое сальдо

⁴⁹ Повышение возрастных коэффициентов рождаемости на 5% означает, что для каждого года прогнозного периода значения базового варианта умножаются на коэффициент 1,05. Например, если в базовом варианте возрастной коэффициент рождаемости для определённой возрастной группы в году t равен $f_{a,t}$, то в сценарии С1 он принимается равным $1,05 \cdot f_{a,t}$. Такая поправка не является ежегодным темпом прироста, а задаёт отклонение сценарной траектории от базовой.

миграции 15 тыс. человек. В сценарии С4 предполагаются снижение возрастных коэффициентов рождаемости на 5 %, ежегодный рост смертности мужчин на 0,3 % и женщин на 0,2 %, а дополнительный миграционный прирост не задаётся.

Параметры сценариев s применяются к базовым рядам рождаемости, смертности и миграции. Сценарные значения демографических компонент рассчитываются по формулам (3.29, 3.30, 3.31):

$$f_{a,t}^{(s)} = k_{fert}^{(s)} \cdot f_{a,t} \quad (3.29)$$

$$m_{a,t}^{(s,g)} = m_{a,t}^{(g)} \cdot \left(1 + r_{mort}^{(s,g)}\right)^{t-t_0} \quad (3.30)$$

$$\mu_{a,t}^{*(s,g)} = \mu_{a,t}^{(g)} + \Delta_{mig}^{(s)} \cdot \omega_{a,g} \quad (3.31)$$

где $f_{a,t}$ обозначает базовый возрастной коэффициент рождаемости для возрастной группы a в году t ; $k_{fert}^{(s)}$ – сценарный коэффициент корректировки рождаемости; $m_{a,t}^{(g)}$ – базовый возрастной коэффициент смертности для пола g ; $r_{mort}^{(s,g)}$ – ежегодное изменение возрастных коэффициентов смертности в сценарии s для пола g ; $\mu_{a,t}^{(g)}$ – базовый возрастно-половой чистый миграционный прирост; $\mu_{a,t}^{*(s,g)}$ – сценарное значение чистого миграционного прироста; $\Delta_{mig}^{(s)}$ – дополнительный годовой чистый миграционный прирост в сценарии s ; $\omega_{a,g}$ – вес распределения дополнительного чистого миграционного прироста по возрастно-половым группам. Индекс s обозначает сценарий, g – пол, a – возрастную группу, t – год прогноза, t_0 – первый год прогнозного периода.

Экзогенные изменения в сценариях задаются в аддитивной форме, поскольку они отражают внешнее отклонение от базового варианта прогноза, а не пропорциональное увеличение уже сложившегося уровня компоненты. Для миграции это особенно важно: чистый миграционный прирост может быть близок к нулю или менять знак, поэтому мультипликативная корректировка теряет содержательный смысл. Аддитивная добавка позволяет задать дополнительный миграционный прирост в абсолютном выражении и распределить его по возрастно-половым группам.

Для каждого сценария было выполнено 10 000 симуляций методом Монте-Карло, в ходе которых моделировалась динамика возрастно-половой структуры населения на ежегодных шагах прогнозирования. Число рождений в году t для сценария s определялось на основе численности женщин репродуктивного возраста по формуле (3.32):

$$B_t^{(s)} = \sum_{a \in A_{repr}} f_{a,t}^{(s)} \cdot P_{a,t}^F, \quad (3.32)$$

где $P_{a,t}^F$ – численность женщин возраста a в момент t . Доля мальчиков при рождении задавалась на основе коэффициента соотношения полов при рождении $SRB = 1,06$, рассчитанного по данным Петростата за 2010-2023 гг., что соответствует $p_m = \frac{SRB}{1+SRB} = 0,514$. Дальнейшая эволюция населения описана когортно-компонентными уравнениями (3.33):

$$\begin{aligned} P_{a+1,t+1}^F &= P_{a,t}^F \times s_{a,t}^F + \mu_{a,t}^{*(s,F)}, & P_{a+1,t+1}^M &= P_{a,t}^M \times s_{a,t}^M + \mu_{a,t}^{*(s,M)}, \\ P_{0,t+1}^F &= B_t^{(s)} \times (1 - p_m), & P_{0,t+1}^M &= B_t^{(s)} \times p_m, \end{aligned} \quad (3.33)$$

где $s_{a,t}$ – вероятность дожития для возраста a . Совокупная численность населения в году t для каждого сценария s определялась как (3.34):

$$N_t^{(s)} = \sum_a (P_{a,t}^F + P_{a,t}^M), \quad (3.34)$$

Затем формировались распределения значений $N_t^{(s)}$ по всем симуляциям, вычислялись медианные оценки и предиктивные интервалы на уровнях 10% и 90% (3.35):

$$\begin{aligned} \text{Median}_t^{(s)} &= \text{median}_d (N_{t,d}^{(s)}), \\ P10_t^{(s)} &= \text{quantile}_{0.10} (N_{t,d}^{(s)}), P90_t^{(s)} = \text{quantile}_{0.90} (N_{t,d}^{(s)}), \end{aligned} \quad (3.35)$$

Программная реализация сценарного модуля выполнена на языке Python и опирается на стохастические модели рождаемости, смертности и миграции, представленные в разделах 3.1.2-3.1.3. Для каждого сценария параметры компонент корректировались в соответствии с заданными сценарными допущениями, после чего выполнен ежегодный когортно-компонентный

погодичный расчет. Такой подход позволяет сохранить общую структуру вероятностной модели и одновременно оценить чувствительность результатов к аналитически заданным изменениям рождаемости, смертности и миграции.

После завершения расчётов для каждого сценария были сформированы траектории численности населения, из которых извлекались медианные значения и предиктивные интервалы. Графическое представление сценарных траекторий и вероятностных диапазонов приведено в разделе 3.2 при интерпретации результатов сценарного анализа.

Сценарный анализ дополняет стохастическую модель, обеспечивая многоаспектную количественную характеристику неопределённости демографического будущего Санкт-Петербурга и позволяя интерпретировать результаты прогноза не только в терминах вероятностей, но и с точки зрения политико-экономических контекстов, определяющих развитие города в долгосрочной перспективе.

3.5. Сценарный анализ демографического развития Санкт-Петербурга (2024-2045 гг.)

После анализа базовой вероятностной траектории был выполнен сценарный расчёт, позволяющий оценить чувствительность демографической динамики Санкт-Петербурга к альтернативным сочетаниям рождаемости, смертности и миграционного прироста. Такая логика сценарного анализа близка к подходам, в которых варианты демографического развития формируются через сочетание гипотез по основным компонентам воспроизводства, включая рождаемость, смертность и миграционный прирост [69]. Были рассчитаны четыре сценария, различающиеся масштабом изменений возрастных коэффициентов рождаемости, смертности и дополнительного чистого миграционного сальдо. В каждом сценарии варьируются три параметра: уровень рождаемости, динамика смертности и добавочное сальдо миграции. Для рождаемости использованы мультипликативные изменения возрастных коэффициентов, что позволяет сохранить возрастной профиль рождаемости и изменить общий уровень деторождения. Для смертности задано ежегодное накопительное изменение возрастных коэффициентов отдельно

для мужчин и женщин, поскольку ожидаемая продолжительность жизни формируется не разовым снижением смертности, а последовательным изменением возрастных рисков на протяжении прогнозного периода. Миграция задавалась как ежегодная аддитивная добавка к чистому миграционному приросту, распределяемая по возрастно-половым группам с концентрацией в возрастах 20-44 года.

Сценарные поправки заданы в масштабе, сопоставимом с фактическими колебаниями демографических показателей Санкт-Петербурга. В прошлые годы суммарный коэффициент рождаемости отклонялся от среднего уровня примерно на 5%, смертность в старших возрастах менялась в близких пределах, а чистый миграционный прирост колебался на десятки тысяч человек в год. Поэтому выбранные параметры не выходят за рамки уже наблюдавшейся изменчивости и используются как реалистичные варианты сценарного расчёта.

Миграционный блок сценариев построен в четырех вариантах дополнительного чистого притока: заданы значения +35 тыс., +25 тыс., +15 тыс. человек в год, а также вариант без дополнительной миграционной добавки. Такой набор вариантов позволяет рассмотреть разные условия восполнения естественной убыли за счёт миграции.

Минимальная положительная добавка +15 тыс. человек в год описывает ситуацию, при которой Санкт-Петербург остаётся привлекательным для переезда, но приток в основном обеспечивается внутренней миграцией. Значение +25 тыс. человек в год соответствует более активному притяжению населения из других регионов России и частично из-за рубежа. В этом случае миграция уже заметнее поддерживает общую численность населения.

Добавка +35 тыс. человек в год задаёт наиболее благоприятный вариант. Он предполагает, что город сохраняет устойчивый приток молодых и трудоспособных групп, связанных с занятостью, образованием и более широкими возможностями крупного рынка труда. Вариант без добавки показывает ситуацию, когда прогноз развивается только по базовой миграционной траектории модели. В совокупности

эти сценарии задают диапазон от сдержанного до усиленного миграционного притока. Параметры четырёх сценариев представлены в таблице 14.

Таблица 14 – Параметры сценариев демографического развития Санкт-Петербурга, 2024-2045 гг.

Сценарий	Изменение возрастных коэффициентов рождаемости, %	Ежегодное изменение возрастных коэффициентов смертности, мужчины, %	Ежегодное изменение возрастных коэффициентов смертности, женщины, %	Дополнительное годовое чистое сальдо миграции, тыс. человек
C1	+5	–1,2	–0,9	+35
C2	+3	–0,8	–0,6	+25
C3	–2	–0,3	–0,2	+15
C4	–5	+0,3	+0,2	0

Примечание. Проценты по рождаемости показывают, на сколько изменяются возрастные коэффициенты рождаемости в каждом сценарии. Проценты по смертности отражают ежегодное накопительное изменение возрастных коэффициентов смертности: отрицательные значения означают ежегодное снижение смертности, положительные – её рост. Итоговые числа родившихся и умерших формируются в ходе когортно-компонентного расчёта с учётом возрастной структуры населения. Поправка к миграции задаётся как ежегодная аддитивная добавка к чистому миграционному приросту, распределяемая по возрастно-половым группам с концентрацией в возрастах 20-44 года.

Более выраженное снижение мужской смертности в благоприятных сценариях интерпретируется как реализация части резерва снижения повышенной смертности мужчин. Для мужчин этот резерв связан прежде всего с причинами, чувствительными к профилактике и изменению поведения: болезнями системы кровообращения, внешними причинами, травматизмом и алкогольно-ассоциированными рисками. Поэтому сценарные параметры смертности отражают не предположение о резком улучшении мужского дожития, а умеренное снижение предотвратимых возрастных рисков на протяжении прогнозного периода.

Ожидаемая продолжительность жизни в работе рассматривается не как внешне заданный сценарный ориентир, а как производный показатель блока смертности вероятностной модели. Такой подход принципиален: прогноз ОПЖ формируется не путём нормативного задания целевых значений, а через динамику возрастных коэффициентов смертности, полученных в результате моделирования.

Тем самым сохраняется внутренняя согласованность прогноза: изменение ОПЖ является следствием трансформации возрастного профиля смертности, а не самостоятельным предположением [118].

Особое внимание было уделено младшим возрастам. Уровень младенческой смертности и её структура, наравне с другими показателями смертности (ожидаемая продолжительность жизни, стандартизованный коэффициент смертности и др.), являются важными индикаторами состояния здоровья населения и социально-экономического развития страны [10]. В исходных данных смертность представлена по укрупнённой группе 0-4 года, тогда как в классической логике построения таблиц смертности младенческий возраст обычно выделяется отдельно. Это связано с тем, что смертность в возрасте до одного года имеет иную природу и интенсивность по сравнению со смертностью детей 1-4 лет. Поэтому для проверки чувствительности расчёта ОПЖ была выполнена условная детализация первой возрастной группы на 0 и 1-4 года.

Данное разделение не является произвольной корректировкой смертности. Для возраста 0 был использован фактический уровень младенческой смертности Санкт-Петербурга в базовом году прогноза – 4,1‰ в 2023 г. Для группы 1-4 года коэффициент смертности рассчитывался остаточным способом, исходя из условия сохранения среднего уровня смертности исходной группы 0-4 года. Такое решение позволяет одновременно учесть специфику младенческой смертности и не нарушить общий уровень смертности, заданный исходными данными. Иными словами, выполняется не изменение режима смертности, а более корректное распределение имеющегося уровня смертности внутри первой пятилетней группы.

Проверка показала, что детализация группы 0-4 не меняет содержательных выводов прогноза, но повышает методическую аккуратность расчёта ОПЖ. Поэтому в основной работе используется вариант с разделением 0 и 1-4 года, расчёт по исходной пятилетней сетке рассматривается как контрольный (табл. 15). Такой подход позволяет сохранить сопоставимость с исходными данными и одновременно приблизить расчёт ОПЖ к международной практике построения кратких таблиц смертности.

Таблица 15 – Ожидаемая продолжительность жизни мужчин и женщин в Санкт-Петербурге при разных вариантах учёта первой возрастной группы, лет

Вариант расчёта	Мужчины, 2035	Женщины, 2035	Мужчины, 2045	Женщины, 2045
Исходная группа 0-4 года	71,19	79,17	70,68	78,6
Детализация 0 и 1-4 года	71,1	79,02	70,6	78,47
Разница, лет	–0,09	–0,15	–0,08	–0,13

Полученные оценки показывают, что блок смертности в модели остаётся устойчивым. Отдельный учёт младшей возрастной группы немного меняет расчётное значение ОПЖ, но не влияет на общий характер результатов. ОПЖ женщин на всём прогнозном горизонте остаётся выше мужской, а разрыв между полами сохраняется. Следовательно, основные выводы определяются не способом задания младших возрастов, а возрастной структурой смертности и её дальнейшей динамикой.

Каждый сценарий задаёт свою комбинацию рождаемости, смертности и миграционного прироста. В сценарии С1 возрастные коэффициенты рождаемости повышаются на 5%, смертность мужчин ежегодно снижается на 1,2%, смертность женщин на 0,9%, а дополнительный миграционный приток составляет 35 тыс. человек в год. Такой вариант можно рассматривать как благоприятный сценарий, при котором улучшение естественного движения населения сочетается с сохранением миграционного притока в молодые и трудоспособные возраста.

Сценарий С2 моделирует промежуточную благоприятную траекторию: рождаемость повышается на 3 %, смертность ежегодно снижается на 0,8 % для мужчин и на 0,6 % для женщин, а дополнительный миграционный прирост составляет 25 тыс. человек в год. Такой сценарий отражает ситуацию, при которой Санкт-Петербург продолжает играть роль центра притяжения населения, но миграционная компенсация выражена слабее, чем в С1.

Сценарий С3 задаёт умеренно неблагоприятную конфигурацию: рождаемость снижается на 2 %, смертность продолжает умеренно снижаться — на 0,3 % в год для мужчин и на 0,2 % для женщин, а дополнительный миграционный прирост сокращается до 15 тыс. человек в год. В этом варианте город ещё сохраняет

миграционное пополнение, однако его масштаб недостаточен для полной компенсации неблагоприятной возрастной инерции.

Наконец, сценарий С4 фиксирует неблагоприятную конфигурацию: рождаемость ниже базовой на 5 %, возрастные коэффициенты смертности ежегодно возрастают на 0,3 % для мужчин и на 0,2 % для женщин, дополнительный миграционный приток отсутствует. Этот сценарий можно трактовать как нижнюю границу рассматриваемого диапазона: это не «крах», но устойчиво неблагоприятная конфигурация, в которой город фактически лишается дополнительного миграционного механизма омоложения и замещения выбывающих поколений. Результаты расчётов представлены на рисунке 19.

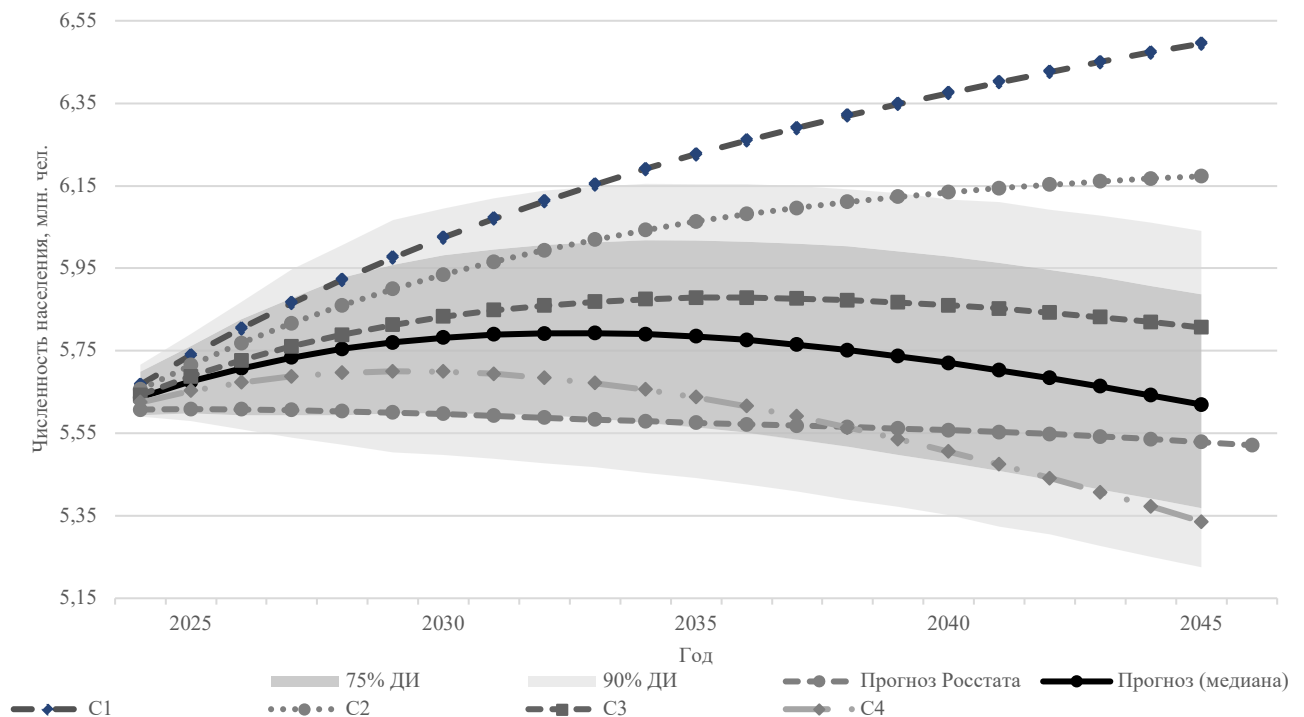


Рисунок 19 – Прогноз численности населения Санкт-Петербурга по сценариям, 2024-2045 гг.

Уже к середине горизонта (середина 2030-х годов) сценарные траектории начинают расходиться: варианты с сохранением повышенного миграционного притока в трудоспособных возрастах ведут к дальнейшему росту численности, тогда как сценарии с ослаблением миграции быстро выходят на стагнацию и последующий спад. К 2045 году диапазон между благоприятным сценарием С1 и

неблагоприятным сценарием С4 составляет около 1,16 млн человек: от 6,495 млн до 5,336 млн. Промежуточные сценарии С2 и С3 формируют более узкий коридор возможных значений – от 5,806 до 6,173 млн человек. Базовая медианная оценка стохастического прогноза, равная 5,620 млн человек, располагается ниже промежуточного коридора и ближе к умеренно неблагоприятной конфигурации, что подчёркивает её осторожный характер и отсутствие заведомо оптимистических допущений по всем компонентам демографического баланса одновременно.

В таблице 16 показаны медианные значения прогнозной численности населения по ключевым годам (2030, 2035, 2045) для каждого сценария. Полные годовые значения медианных траекторий сценарного прогноза представлены в приложении 10 (табл. 1).

Таблица 16 – Прогнозная численность населения Санкт-Петербурга по сценариям, тыс. человек

Сценарий	2030	2035	2045
С1	6 024,50	6 226,60	6 494,60
С2	5 934,60	6 063,50	6 172,80
С3	5 832,70	5 878,50	5 806,40
С4	5 699,20	5 637,60	5 335,80

Источник: рассчитано автором.

Ключевой вывод из таблицы 16 состоит в постепенном расхождении сценарных траекторий на прогнозном горизонте. В 2030 г. различия между сценариями ещё относительно умеренны, поскольку краткосрочная динамика в значительной степени определяется уже сложившейся возрастной структурой населения. К 2035 г. разрыв усиливается, а к 2045 г. расхождение между С1 и С4 превышает 1,1 млн человек. Основная причина этого расхождения – накопленный миграционный эффект: при заданных изменениях рождаемости и смертности именно миграция остаётся главным фактором, раздвигающим траектории численности населения мегаполиса.

Сценарии различаются не только итоговой численностью населения, но и глубиной перестройки возрастной структуры. Во всех вариантах сохраняется движение к старению возрастного профиля, однако его выраженность неодинакова. В сценарии С1 рост общей численности и более благоприятный режим дожития

смягчают относительное старение населения: к 2045 г. доля жителей 65 лет и старше составляет около 19,8%. В неблагоприятном сценарии С4 старение выражено сильнее: доля населения 65 лет и старше достигает 24,1%, поскольку сокращение младших когорт сочетается с недостаточной миграционной компенсацией и динамикой смертности. Сценарные значения ожидаемой продолжительности жизни при рождении по полу представлены в приложении 10 (табл. 2).

Наиболее заметные изменения проявляются в старших возрастах. Численность населения 80 лет и старше к 2045 г. увеличивается примерно на 77%, а группа 85 лет и старше почти удваивается. Одновременно численность детей 0-14 лет сокращается примерно на 16 п.п. Следовательно, различия между сценариями нельзя сводить только к вопросу роста или снижения общей численности: ключевая развилка состоит в том, насколько глубокой окажется перестройка возрастной структуры и сможет ли миграционный приток смягчить последствия старения.

Вместе с тем, расхождение между официальной и медианной в вероятностном прогнозе не является принципиальным: официальный прогноз на всём горизонте остаётся в рамках 90%-го доверительного интервала стохастической модели. К концу периода различия сокращаются: к 2045 г. Росстат оценивает численность населения Санкт-Петербурга на уровне 5,52 млн человек, тогда как медианная траектория модели показывает 5,62 млн человек. Таким образом, официальный прогноз оказывается ниже медианной оценки, но укладывается в область вероятных траекторий, задаваемую стохастической моделью.

Сценарные кривые (рис. 19) показывают, что основной разброс долгосрочных траекторий связан не с начальным уровнем численности, а с различиями в предпосылках по миграции и естественному движению населения. Благоприятные сценарии формируют устойчивый рост численности, тогда как сценарии с ослаблением миграционной компенсации приводят к снижению населения к нижней части прогнозного диапазона.

В результате сценарный блок показывает, что устойчивость численности населения Санкт-Петербурга на долгосрочном горизонте определяется, прежде всего масштабом миграционного притока лиц трудоспособного и предпенсионного возраста. Значимость трудоспособной возрастной структуры миграционного притока согласуется с результатами исследований, где миграция рассматривается как фактор, влияющий не только на общую численность населения, но и на динамику состава рабочей силы [5]. Даже при умеренных изменениях рождаемости и смертности город способен удерживать численность жителей более 6 млн человек к середине 2040-х годов, если годовой чистый миграционный прирост будет на уровне 30-40 тыс. человек. При исчерпании миграционного ресурса (нулевая дополнительная миграция) численность населения к концу периода прогнозирования снизится и приблизится к нижней границе предиктивного интервала базового прогноза.

Выводы к 3 главе. К 2045 г. Санкт-Петербург сохранит масштаб второго по величине мегаполиса, но войдёт в новый возрастно-структурный режим. Медианная траектория базового прогноза показывает сохранение численности населения вблизи уровня 2023 г., но за этой устойчивостью будут стоять старение населения и высокая зависимость города от миграции. Прогноз показывает, что главная демографическая проблема Санкт-Петербурга будет связана не только с будущей численностью населения. Более существенным станет изменение внутреннего состава города. Сужение младших возрастов, рост численности старших когорт и постепенное сокращение доли населения трудоспособных возрастов сформируют новую нагрузку на рынок труда, социальную инфраструктуру, здравоохранение и систему стратегического управления городским развитием.

Вероятностная когортно-компонентная модель позволила представить это будущее не одной расчётной линией, а системой возможных траекторий. Стохастическое моделирование рождаемости, смертности и миграции, Монте-Карло-симуляции и калибровка предиктивных интервалов сохраняют логику классического компонентного расчёта, но снимают его главный недостаток –

привязку к одному сценарию. Сопоставление с прогнозом Росстата показывает близость общего направления оценок, однако стохастическая модель дополнительно раскрывает диапазон неопределённости и показывает, какие компоненты формируют этот разброс.

Сценарный анализ уточняет содержательный смысл прогноза. Различия между траекториями будут определяться прежде всего масштабом миграционной компенсации. При устойчивом притоке населения трудоспособных возрастов город сможет смягчать последствия старения и поддерживать рост. При ослаблении миграционного притяжения старение станет более выраженным, а численность населения сместится к нижней границе прогнозного диапазона. Рождаемость и смертность задают важные ограничения развития, но именно миграция останется компонентой, от которой будет зависеть глубина долгосрочной демографической развилки.

Вероятностный прогноз численности населения Санкт-Петербурга до 2045 г. показывает, что для мегаполиса управленчески значима не только итоговая оценка населения, но и условия, при которых эта оценка будет достигаться. Предложенный подход позволяет рассматривать демографическое будущее города как сочетание инерции возрастной структуры, изменчивости миграционного режима и неопределённости компонент воспроизводства. Поэтому вероятностная модель выступает как навигатор статистической поддержки стратегических решений развития мегаполиса.

ЗАКЛЮЧЕНИЕ

Санкт-Петербург, как мегаполис и ядро Северо-Западного федерального округа, обоснованно рассматривается как открытая демографическая система. В отличие от национального уровня, где потоки прибывших и выбывших внутренней миграции взаимно уравниваются, для Санкт-Петербурга миграция выступает самостоятельным фактором динамики численности и возрастно-половой структуры населения.

Проведённое исследование показало, что демографическое развитие Санкт-Петербурга имеет выраженную компонентную асимметрию. Рождаемость сохраняет профиль, свойственный крупным городам мира: позднее материнство, низкую интенсивность рождений и ограниченный вклад в естественное воспроизводство. На уровне смертности всё заметнее сказываются сдвиги в возрастной структуре. Увеличение численности и доли старших когорт поддерживает высокий абсолютный уровень числа умерших даже при снижении возрастных коэффициентов. Миграция остаётся главным механизмом компенсации естественной убыли, но её воздействие неоднородно по возрасту и полу, экономически чувствительно и не полностью отражается в регистрационной статистике.

С учётом этих особенностей в диссертации обоснован и применён вероятностный когортно-компонентный подход к прогнозированию численности и возрастно-половой структуры населения Санкт-Петербурга до 2045 г. Применение этого подхода исходит из того, что мегаполис не является уменьшенной копией страны или обычного региона: его демографическая динамика формируется в условиях высокой открытости, интенсивного миграционного обмена, неполной наблюдаемости мобильного населения и повышенной чувствительности к внешним изменениям. Поэтому прогноз численности населения для такого мегаполиса как Санкт-Петербург, должен быть представлен не одной расчётной траекторией, а рядом возможных исходов, возникающих при различной динамике рождаемости, смертности и миграции.

Методический результат работы состоит в адаптации к условиям мегаполиса и применении вероятностной схемы прогноза на базе официальной статистики, приведённой к единой возрастно-половой структуре. На разных этапах исследования были использованы методы вторичной группировки исходных данных, стохастическое моделирование возрастных профилей рождаемости, смертности и миграции методом PCA-AR(1), Монте-Карло-симуляции и калибровка предиктивных интервалов по результатам ретроспективной проверки. Такая постановка сохраняет демографическую природу когортно-компонентного расчёта и одновременно устраняет однозначность фиксированного сценария.

Базовый прогноз показал сохранение численности населения Санкт-Петербурга (при некотором уменьшении) при продолжающейся перестройке возрастной структуры – старении населения. По медианной траектории численность населения к 2045 г. составит 5,62 млн человек. Официальный прогноз Росстата находится в рамках предиктивного диапазона авторской модели, приближаясь к нижней границе 90%-го доверительного интервала, что подтверждает сопоставимость результатов расчетов с официальной оценкой. Различие состоит в аналитической постановке: официальный прогноз фиксирует одну траекторию, тогда как вероятностная модель показывает диапазон возможных отклонений и выявляет компоненты, определяющие этот разброс.

Сценарные расчёты подтвердили, что долгосрочная развилка демографического развития Санкт-Петербурга определяется масштабом миграционной компенсации. При благоприятной конфигурации рождаемости, смертности и миграции численность населения к 2045 г. достигает 6,495 млн человек; при неблагоприятной конфигурации снижается до 5,336 млн человек. Этот разрыв формируется прежде всего накопленным эффектом миграции трудоспособных возрастов. Естественное движение в рассматриваемом горизонте не становится самостоятельным источником устойчивого роста.

Главный содержательный вывод исследования состоит в том, что демографическая устойчивость Санкт-Петербурга не может оцениваться только по итоговой численности населения. При близких значениях общей численности

возможны разные возрастные структуры, разные масштабы старения и разная нагрузка на трудоспособное население. Для мегаполиса миграция выступает не остаточным дополнением к естественному приросту, а структурным компонентом воспроизводства, одновременно меняющим численность, возрастной состав и потенциал дальнейшего развития города.

Практическое значение применённого подхода, адаптированного к условиям мегаполиса заключается в возможности использования прогноза как инструмента оценки демографических рисков. Для городского управления важна не одна средняя оценка, а диапазон траекторий, в пределах которого должны разрабатываться и проверяться решения в сфере жилья, транспорта, здравоохранения, образования, рынка труда и системы услуг. Значение таких оценок возрастает в условиях структурных изменений занятости: М. А. Клупт, О. С. Коротева, А. В. Курячий и Д. С. Чернейко рассматривают «структурные дисбалансы рынка труда» как сочетание кадрового дефицита в одних профессиях и избытка персонала или сложностей трудоустройства в других [25]. Поэтому прогноз возрастной структуры населения мегаполиса может использоваться не только для оценки общей демографической нагрузки, но и для информационной поддержки решений в сфере занятости и профессионального образования. Для управленческих задач особенно важна детализация информации о старших возрастах. Примером специализированного информационного ресурса является зарегистрированная база данных «Медико-социальный статус населения 80-ти лет и старше», включающая показатели половозрастной структуры, семейного положения, социально-экономического статуса и хронических заболеваний лиц 80 лет и старше [58].

Санкт-Петербург вступает в период, когда устойчивость будет определяться не просто приростом населения, а сохранением трудоспособного ядра, управляемостью миграционного притока и способностью смягчать последствия старения. Поэтому вероятностный когортно-компонентный прогноз является для мегаполиса не техническим усложнением расчёта, а более точной постановкой задачи в навигации демографического прогнозирования.

СПИСОК ЛИТЕРАТУРЫ

1. Андреев, Е. М. Смертность в Москве и других мегаполисах мира: сходства и различия / Е. М. Андреев, Е. А. Кваша, Т. Л. Харькова // Демографическое обозрение. – 2016. – Т. 3, № 3. – С. 39-79. – EDN WZZJGH.
2. Араб-Оглы, Э. А. Демографические и экологические прогнозы. Критика современных буржуазных концепций / Э. А. Араб-Оглы. – Москва : Статистика, 1978. – 319 с.
3. Боченина, М. В. Эффективность региональной системы здравоохранения: рейтинг субъектов Российской Федерации / М. В. Боченина, С. М. Кургина, С. А. Музалевская, Т. А. Янгиров // Экономика и управление: научно-практический журнал. – 2025. – № 1(181). – С. 29-35. – DOI 10.34773/EU.2025.1.5.
4. Бугаев, М. А. Маятниковые миграции на рынке труда Санкт-Петербурга и Ленинградской области / М. А. Бугаев // Вестник Санкт-Петербургского университета. Экономика. – 2014. – № 4. – С. 86-116.
5. Варшавская, Е. Я. Миграция и динамика численности и состава рабочей силы в России до 2050 г. / Е. Я. Варшавская, М. Б. Денисенко // Население постсоветских стран: тенденции, прогнозы, межпоколенческие трансферты : монография / под науч. ред. М. Б. Денисенко. – Москва : Прометей, 2024. – С. 153-186.
6. Василенко, П. В. Применение гравитационной модели для анализа внутриобластных миграций на примере Новгородской и Псковской областей / П. В. Василенко // Псковский регионологический журнал. – 2013. – № 15. – С. 83-90.
7. Вишневский, А. Г. Демографические тревоги России / А. Г. Вишневский, С. В. Захаров // Вестник АН СССР. – 1990. – № 8. – С. 16-33.
8. Гагауз, О. Е. Разработка демографических прогнозов в Молдове: опыт прогнозирования и результаты / О. Е. Гагауз, И. В. Пахомий, М. Слав // Население постсоветских стран: тенденции, прогнозы, межпоколенческие трансферты : монография / под науч. ред. М. Б. Денисенко. – Москва : Прометей, 2024. – С. 26-74.

9. Демографический прогноз России до 2046 года : презентационные материалы / Федеральная служба государственной статистики. – Москва : Росстат, 2023. – URL: <https://rosstat.gov.ru/folder/12781>
10. Денисенко, М. Б. Младенческая смертность в Москве в 1939-1949 гг. / М. Б. Денисенко, Е. А. Кваша, Ю. В. Хавраева // Демографическое обозрение. – 2025. – Т. 12, № 1. – С. 57-85. – DOI 10.17323/demreview.v12i1.26577.
11. Елисеева, И. И. Изменения в структуре российских домохозяйств/семей: возможный сценарий / И. И. Елисеева, М. П. Декина // Статистика и Экономика. – 2025. – Т. 22, № 6. – С. 9-18. – DOI 10.21686/2500-3925-2025-6-9-18.
12. Елисеева, И. И. Основная семья: предыстория и перспективы / И. И. Елисеева // Демографическое обозрение. – 2020. – Т. 7, № 4. – С. 6-35. – DOI 10.17323/demreview.v7i4.12042. – EDN LAPCWW.
13. Елисеева, И. И. Перспективные оценки численности населения Санкт-Петербурга / И. И. Елисеева, М. В. Боченина // Санкт-Петербург: глобальный город. – Санкт-Петербург, 2018.
14. Елисеева, И. И. Рождения и нерождения / И. И. Елисеева // Петербургская социология сегодня. – 2018. – № 10. – С. 204-216. – DOI 10.25990/socinstras.pss-10.g1cq-yc02. – EDN VRWANG.
15. Захаров, С. В. История рождаемости в России: от поколения к поколению / С. В. Захаров // Демографическое обозрение. – 2023. – Т. 10, № 1. – С. 4-43. – DOI 10.17323/demreview.v10i1.17259. – EDN STMVEB.
16. Калабихина И. Е. «Гендерный переход» и демографическое развитие // Электронный журнал. – 2009.
17. Карачурина, Л. Б. Внутрорегиональная миграция населения в России: пригороды выигрывают у столиц / Л. Б. Карачурина, Н. В. Мкртчян // Известия Российской академии наук. Серия географическая. – 2021. – Т. 85, № 1. – С. 24-38. – DOI 10.31857/S2587556621010076.
18. Клупт, М. А. Влияние семейной политики и нормативных представлений о семье на рождаемость: компаративный анализ / М. А. Клупт // Социологические исследования. – 2020. – № 3. – С. 40-50. – DOI 10.31857/S013216250008812-6.

19. Клулт, М. А. Демографическое развитие в свете кризисов XXI века: теоретический анализ / М. А. Клулт // Социологические исследования. – 2022. – № 11. – С. 15-24. – DOI 10.31857/S013216250022083-4.
20. Клулт, М. А. Демография регионов Земли / М. А. Клулт. – Санкт-Петербург : Питер, 2008. – 347 с.
21. Клулт, М. А. Запад, не-Запад и стратегии политической борьбы на демографическом поле / М. А. Клулт // Общественные науки и современность. – 2022. – № 1. – С. 78-90. – DOI 10.31857/S0869049922010051.
22. Клулт, М. А. Интеграция данных о населении: проблемы и перспективы / М. А. Клулт, О. Н. Никифоров // Вопросы статистики. – 2021. – Т. 28, № 3. – С. 5-14. – DOI 10.34023/2313-6383-2021-28-3-5-14.
23. Клулт, М. А. Парадигмы и оппозиции современной демографии / М. А. Клулт // Демографическое обозрение. – 2014. – Т. 1, № 1. – С. 34-56.
24. Клулт, М. А. Проблемы семьи и рождаемости в ценностных конфликтах 2010-х гг. / М. А. Клулт // Социологические исследования. – 2021. – № 5. – С. 36-46. – DOI 10.31857/S013216250014119-3.
25. Клулт, М. А. Структурные дисбалансы рынка труда: анализ через призму профессий / М. А. Клулт, О. С. Коротеева, А. В. Курячий, Д. С. Чернейко // Диалог. – 2024. – № 3. – С. 58-71.
26. Клулт, М. А. Тревоги XXI века: механизмы влияния на рождаемость / М. А. Клулт // Социологические исследования. – 2022. – № 5. – С. 25-35. – DOI 10.31857/S013216250018464-3.
27. Коломак, Е. А. Развитие городской системы России: тенденции и факторы / Е. А. Коломак // Вопросы экономики. – 2014. – № 10. – С. 82-96. – EDN SVUXRZ.
28. Конференция «Петербург — Барселона»: возможности гармоничного развития городов [Электронный ресурс] // Университет ИТМО. – URL: <https://news.itmo.ru/ru/news/5763/> (дата обращения: 16.02.2024).
29. Кубасова, Е. И. Анализ основных тенденций рождаемости населения двух столиц / Е. И. Кубасова // Актуальные проблемы учета, анализа, финансового контроля и статистики в социальной сфере : сборник Межвузовской конференции

научных работ студентов, аспирантов и профессорско-преподавательского состава, Москва, 07 апреля 2017 года / под ред. В. А. Свободина, С. В. Шамшеева. – Москва : Научный консультант, 2017. – С. 48-55. – EDN YRUDLT.

30. Кубасова, Е. И. Оценка уровня рождаемости населения Москвы и Санкт-Петербурга / Е. И. Кубасова // Актуальные проблемы учета, анализа, финансового контроля и статистики в социальной сфере : сборник Межвузовской конференции научных работ студентов, аспирантов и профессорско-преподавательского состава, Москва, 07 апреля 2017 года / под ред. В. А. Свободина, С. В. Шамшеева. – Москва : Научный консультант, 2017. – С. 148-156. – EDN YRUDSH.

31. Кубасова, Е. И. Сравнительный анализ дифференциации рождаемости населения Москвы и Санкт-Петербурга / Е. И. Кубасова // Актуальные проблемы учета, анализа и аудита в социальной сфере : сборник научных работ студентов, аспирантов и профессорско-преподавательского состава, Москва, 06 апреля 2016 года. – Москва : Научный консультант, 2016. – С. 148-155. – EDN VXJLVJ.

32. Лексин, В. Н. Мегаполисы и феномен мегаполисности в России / В. Н. Лексин, Б. Н. Порфирьев // Регион: экономика и социология. – 2017. – № 1(93). – С. 42-77. – DOI 10.15372/REG20170103. – EDN VVEZOK.

33. Локосов, В. В. Демографическое развитие России: динамика и социально-экономические риски / В. В. Локосов // Вестник Российской академии наук. – 2020. – Т. 90, № 3. – С. 251-258. – DOI 10.31857/S086958732003010X. – EDN BOPSEX.

34. Ломоносов, М. В. О сохранении и размножении российского народа / М. В. Ломоносов // Пространство и Время. – 2011. – № 4(6). – С. 113-119. – EDN OORUKB.

35. Менделеев, Д. И. К познанию России / Д. И. Менделеев. – 7-е изд. – Санкт-Петербург : Издание А. С. Суворина, 1912. – 160 с. : табл., карты.

36. Мкртчян, Н. В. Крупные города России и их пригороды как центры притяжения внутренних мигрантов / Н. В. Мкртчян, Р. И. Гильманов // Вестник Санкт-Петербургского университета. Науки о Земле. – 2023. – Т. 68, № 1. – С. 44-63. – DOI 10.21638/spbu07.2023.103. – EDN LXYZLA.

37. Мкртчян, Н. В. Внутренняя миграция в России в 2010-е гг.: макрорегиональные особенности / Н. В. Мкртчян // Демографическое обозрение. – 2023. – Т. 10, № 3. – С. 21-42. – DOI 10.17323/demreview.v10i3.17968. – EDN LOZTGU.
38. Мкртчян, Н. В. Внутренняя миграция как ресурс развития России: социально-экономические эффекты, издержки и ограничения : научный доклад / Н. В. Мкртчян, Ю. Ф. Флоринская, К. И. Казенин. – Москва : Дело РАНХиГС, 2020. – 76 с. – ISBN 978-5-85006-190-6.
39. Мкртчян, Н. В. Трудовая миграция в России: международный и внутренний аспекты / Н. В. Мкртчян, Ю. Ф. Флоринская // Журнал Новой экономической ассоциации. – 2018. – № 1(37). – С. 186-193. – DOI 10.31737/2221-2264-2018-37-1-9. – EDN XRWLZR.
40. Население и глобализация : монография / Н. М. Римашевская, В. Ф. Галецкий, А. А. Овсянников [и др.] ; под общ. ред. Н. М. Римашевской ; Российская академия наук, Институт социально-экономических проблем народонаселения. – Москва : Наука, 2002. – 321, [1] с. – ISBN 5-02-006227-8.
41. Население постсоветских стран: тенденции, прогнозы, межпоколенческие трансферты / науч. ред. М. Б. Денисенко. – М. : Прометей, 2024.
42. Настоящее и будущее семьи в меняющемся мире : коллективная монография / под науч. ред. Н. М. Римашевской, В. Г. Доброхлеб ; Институт социально-экономических проблем народонаселения РАН. – Т. 2. – Москва : Экономическое образование, 2018. – 216 с. – ISBN 978-5-7425-0189-3. – DOI 10.26653/2017/978-5-7425-0189-3. – EDN AFGHLB.
43. Невский — Рамбла: встречное движение. [Электронный ресурс]. – URL: <http://img.rosbalt.ru/doc/style/nevsky-rambla.pdf> (дата обращения: 16.02.2024).
44. Нерадовская, Ю. В. Оценка эффективности функционирования региональных систем здравоохранения / Ю. В. Нерадовская // Экономика и управление: научно-практический журнал. – 2022. – № 1(163). – С. 125-132. – DOI 10.34773/EU.2022.1.24.

45. Ниворожкина Л. И., Ниворожкин А. М., Арженовский С. В. «Цена» материнства: эконометрическая оценка. – 2008.
46. Никитина, С. Ю. Вероятностный прогноз численности населения России / С. Ю. Никитина, С. Я. Щербов // Вопросы статистики. – 2007. – № 7. – С. 6-9.
47. Никитина, С. Ю. Совершенствование статистической методологии прогноза численности населения в условиях недостатка демографических данных : автореф. дис. ... канд. экон. наук : 08.00.12 / С. Ю. Никитина. – Москва, 2008. – 12 с.
48. Носова, М. Г. Стохастическая модель для прогнозирования численности населения / М. Г. Носова // Бюллетень науки и практики. – 2019. – Т. 5, № 9. – С. 17-25. – DOI 10.33619/2414-2948/46/02.
49. Порфирьев, Б. Н. Города и мегаполисы: проблема дефиниций и индикаторы устойчивого развития / Б. Н. Порфирьев, С. Н. Бобылёв // Проблемы прогнозирования. – 2018. – № 2 (167). – С. 14-23. – URL: <https://ecfor.ru/wp-content/uploads/2018/06/indikatory-razvitiya-gorodov.pdf>
50. Римашевская, Н. М. Основные направления совершенствования демографической и семейной политики в современной России / Н. М. Римашевская, В. Г. Доброхлеб // Народонаселение. – 2013. – № 1(59). – С. 30-41. – EDN QABSDb.
51. Рыбаковский, Л. Л. Депопуляция в России: итоги за 1992-2022 гг., компоненты, компенсация миграцией на региональном уровне / Л. Л. Рыбаковский, О. Л. Рыбаковский // Социально-трудовые исследования. – 2023. – № 2(51). – С. 16-26. – DOI 10.34022/2658-3712-2023-51-2-16-26. – EDN GZOCNT.
52. Рыбаковский, Л. Л. Методологические вопросы прогнозирования населения / Л. Л. Рыбаковский. – Москва : Статистика, 1978. – 208 с.
53. Рябчикова, Д. А. Анализ динамики внутренней миграции в г. Санкт-Петербурге с использованием гравитационной модели / Д. А. Рябчикова // Вопросы статистики. – 2025. – Т. 32, № 2. – С. 40-51. – DOI 10.34023/2313-6383-2025-32-2-40-51. – EDN UAWKJ.
54. Рябчикова, Д. А. Прогнозирование численности населения современного мегаполиса (на примере Санкт-Петербурга): методологические аспекты / Д. А.

Рябчикова // Экономический журнал ВШЭ. – 2026. – Т. 30, № 1. – С. 182-205. – DOI 10.17323/ej.2026.33626.

55. Рябчикова, Д. А. Смертность в двух крупнейших мегаполисах России / Д. А. Рябчикова // Народонаселение. – 2025. – Т. 28, № 2. – С. 29-41. – DOI 10.24412/1561-7785-2025-2-29-41. – EDN WDFTGF.

56. Сафарова Г. Л. Демографические закономерности старения населения : дис. ... д-ра биол. наук. – Санкт-Петербург, 2002.

57. Сафарова Г. Л. Демография старения: современное состояние и приоритетные направления исследований // Успехи геронтологии. – 2009.

58. Свидетельство о государственной регистрации базы данных № 2021621415 Российская Федерация. Медико-социальный статус населения 80-ти лет и старше : № 2021620447 : заявл. 17.03.2021 : опубл. 01.07.2021 / М. Л. Сиротко, М. Б. Денисенко, И. А. Золотовская, М. В. Комарова. – EDN HTTYCL.

59. Синицына, Д. А. Рождаемость в двух столичных городах России / Д. А. Синицына // Народонаселение. – 2024. – Т. 27, № 2. – С. 125-137. – DOI 10.24412/1561-7785-2024-2-129. – EDN QOLQJW.

60. Сороко Е. Л. О пределах применимости демографических прогнозов ООН // Демографическое обозрение. – 2018. – Т. 5, № 2. – С. 6-31. – DOI 10.17323/demreview.v5i2.7933.

61. Струмилин, С. Г. Наши трудовые ресурсы и перспективы / С. Г. Струмилин // Хозяйственное строительство. – 1922. – Вып. 2. – С. 32-46.

62. Урланис, Б. Ц. Проблемы динамики населения СССР / Б. Ц. Урланис. – Москва : Наука, 1974. – 335 с.

63. Федорищенко, В. М. Влияние миграции на население Санкт-Петербурга / В. М. Федорищенко // Скиф. Вопросы студенческой науки. – 2020. – № 12(52). – С. 70-77. – EDN ННШВЕ.

64. Харькова, Т. Л. Сравнительная оценка смертности населения в российских и зарубежных мегаполисах / Т. Л. Харькова, Е. А. Кваша, Б. А. Ревич // Проблемы прогнозирования. – 2018. – № 6(171). – С. 150-159. – EDN YYJJXV.

65. Чистякова, Н. Е. Миграция в Санкт-Петербурге: демографический аспект / Н. Е. Чистякова // Социология вчера, сегодня, завтра : Вторые социологические чтения памяти Валерия Борисовича Голофаства / под ред. О. Б. Божкова. – Санкт-Петербург : Эйдос, 2008. – С. 120-128. – EDN UAWDGD.
66. Чудиновских, О. С. О качестве федерального статистического наблюдения за миграционными процессами / О. С. Чудиновских, А. В. Степанова // Демографическое обозрение. – 2020. – Т. 7, № 1. – С. 54-82. – DOI 10.17323/demreview.v7i1.10820. – EDN CGSRIL.
67. Чудиновских, О. С. О новых технологиях и статистике миграции в России / О. С. Чудиновских, Е. В. Донец // Вопросы статистики. – 2018. – Т. 25, № 5. – С. 11-26.
68. Чудиновских, О. С. О понимании статистики миграции / О. С. Чудиновских // Вопросы статистики. – 2017. – № 5. – С. 19-27.
69. Юмагузин, В. В. Прогноз численности и демографической нагрузки населения России до 2100 года / В. В. Юмагузин, М. В. Винник // Проблемы прогнозирования. – 2022. – № 4. – С. 98-111. – DOI 10.47711/0868-6351-193-98-111.
70. Abel, G. J. Bilateral international migration flow estimates for 200 countries / G. J. Abel, J. E. Cohen // Scientific Data. – 2019. – Vol. 6. – Article 82. – P. 1-13. – DOI 10.1038/s41597-019-0089-3.
71. Alexander, M. A Bayesian cohort component projection model to estimate women of reproductive age at the subnational level in data-sparse settings / M. Alexander, L. Alkema // Demography. – 2022. – Vol. 59, № 5. – P. 1713-1737. – DOI 10.1215/00703370-10216406.
72. Alho J. M. Stochastic methods in population forecasting // Int. J. Forecasting. 1990. Vol. 6, № 4. P. 521-530. DOI: 10.1016/0169-2070(90)90030-F.
73. Alho, J. M. Statistical Demography and Forecasting / J. M. Alho, B. D. Spencer. – New York : Springer, 2005. – 410 p.
74. Azose, J. J. Bayesian probabilistic projection of international migration / J. J. Azose, A. E. Raftery // Demography. – 2015. – Vol. 52, № 5. – P. 1627-1650. – DOI 10.1007/s13524-015-0415-0.

75. Azose, J. J. Estimation of emigration, return migration, and transit migration between all pairs of countries / J. J. Azose, A. E. Raftery // *Proceedings of the National Academy of Sciences*. – 2019. – Vol. 116, № 1. – P. 116-122. – DOI 10.1073/pnas.1722334116.
76. Bijak, J. Bayesian forecasting of immigration to selected European countries by using expert knowledge / J. Bijak, A. Wiśniowski // *Journal of the Royal Statistical Society: Series A*. – 2010. – Vol. 173, № 4. – P. 775-796. – DOI 10.1111/j.1467-985X.2009.00635.x.
77. Bongaarts, J. On the quantum and tempo of fertility / J. Bongaarts, G. Feeney // *Population and Development Review*. – 1998. – Vol. 24, № 2. – P. 271-291. – DOI 10.2307/2807974.
78. Booth, H. Demographic forecasting: 1980 to 2005 in review / H. Booth // *International Journal of Forecasting*. – 2006. – Vol. 22, № 3. – P. 547-581. – DOI 10.1016/j.ijforecast.2006.04.001.
79. Booth, H. Mortality modelling and forecasting: A review of methods / H. Booth, L. Tickle // *Annals of Actuarial Science*. – 2008. – Vol. 3, № 1-2. – P. 3-43. – DOI 10.1017/S1748499500000440.
80. Box, G. E. P. *Time Series Analysis: Forecasting and Control* / G. E. P. Box, G. M. Jenkins, G. C. Reinsel, G. M. Ljung. – 5th ed. – Hoboken : Wiley, 2015. – 712 p. – ISBN 978-1-118-67502-1.
81. Cameron, M. Lessons from Stochastic Small-Area Population Projections: The Case of Waikato Subgroups in New Zealand / M. Cameron, J. Poot // *Journal of Population Research*. – 2011. – Vol. 28. – P. 245-265. – DOI 10.1007/s12546-011-9056-3.
82. *Cities in the World: A New Perspective on Urbanisation* / OECD, European Commission. – Paris : OECD Publishing, 2020. – DOI 10.1787/d0efcbda-en.
83. Cottineau, C. Defining urban clusters to detect agglomeration economies / C. Cottineau, O. Finance, E. Hatna, E. Arcaute, M. Batty // *Environment and Planning B: Urban Analytics and City Science*. – 2019. – Vol. 46, № 9. – P. 1611-1626. – DOI 10.1177/2399808318755146.

84. Dyrting, S. Smoothing migration intensities with P-TOPALS / S. Dyrting // Demographic Research. – 2020. – Vol. 43, Article 55. – P. 1607-1650. – DOI 10.4054/DemRes.2020.43.55.
85. Greenwood, M. J. Human migration: theory, models, and empirical studies / M. J. Greenwood // Journal of Regional Science. – 1985. – Vol. 25, № 4. – P. 521-544. – DOI 10.1111/j.1467-9787.1985.tb00321.x.
86. Hyndman, R. J. Forecasting: principles and practice / R. J. Hyndman, G. Athanasopoulos. – 3rd ed. – Melbourne : OTexts, 2021. – URL: <https://otexts.com/fpp3/>
87. Hyndman, R. J. Mortality forecasting: a review of methods / R. J. Hyndman, R. Lee, L. Tickle // Statistical Modelling. – 2003. – Vol. 3, № 4. – P. 289-303.
88. Hyndman, R. J. Robust forecasting of mortality and fertility rates: A functional data approach / R. J. Hyndman, M. S. Ullah // Computational Statistics & Data Analysis. – 2007. – Vol. 51, № 10. – P. 4942-4956. – DOI 10.1016/j.csda.2006.07.028.
89. Hyndman, R. J. Stochastic population forecasts using functional data models for mortality, fertility and migration / R. J. Hyndman, H. Booth // International Journal of Forecasting. – 2008. – Vol. 24, № 3. – P. 323-342. – DOI 10.1016/j.ijforecast.2008.02.009.
90. The intricate interplay between cities and migration / United Nations Department of Economic and Social Affairs. – 2017. – 30 August. – URL: <https://www.un.org/development/desa/en/news/population/cities-and-migration.html>
91. Keilman, N. Why population forecasts should be probabilistic – illustrated by the case of Norway / N. Keilman, D. Q. Pham, A. Hetland // Demographic Research. – 2002. – Vol. 6. – P. 409-454. – DOI 10.4054/DemRes.2002.6.15.
92. Keyfitz, N. The limits of population forecasting / N. Keyfitz // Population and Development Review. – 1981. – Vol. 7, № 4. – P. 579-593. – DOI 10.2307/1972799.
93. Lee, R. D. Modeling and forecasting U.S. mortality / R. D. Lee, L. R. Carter // Journal of the American Statistical Association. – 1992. – Vol. 87, № 419. – P. 659-671.
94. Lee, R. Evaluating the performance of the Lee-Carter method for forecasting mortality / R. Lee, T. Miller // Demography. – 2001. – Vol. 38, № 4. – P. 537-549. – DOI 10.1353/dem.2001.0036.

95. Lesthaeghe, R. The unfolding story of the second demographic transition / R. Lesthaeghe // *Population and Development Review*. – 2010. – Vol. 36, № 2. – P. 211-251. – DOI 10.1111/j.1728-4457.2010.00328.x.
96. Li, N. Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method / N. Li, R. Lee // *Demography*. – 2005. – Vol. 42, № 3. – P. 575-594. – DOI 10.1353/dem.2005.0021.
97. Lutz, W. Probabilistic population projections based on expert opinion / W. Lutz, W. Sanderson, S. Scherbov // *Nature*. – 1998. – Vol. 397. – P. 710-713.
98. Methodological Manual on City Statistics : 2017 edition / Eurostat. – Luxembourg : Publications Office of the European Union, 2017. – DOI 10.2785/708009. – URL: <https://ec.europa.eu/eurostat/documents/3859598/8012444/KS-GQ-17-006-EN-N.pdf>
99. Population projections in the EU – methodology // Eurostat Statistics Explained. – 2026. – URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Population_projections_in_the_EU_-_methodology (дата обращения: 16.05.2026).
100. Raftery, A. E. Bayesian probabilistic population projections for all countries / A. E. Raftery, N. Li, H. Ševčíková, P. Gerland, G. K. Heilig // *Proceedings of the National Academy of Sciences*. – 2012. – Vol. 109, № 35. – P. 13915-13921. – DOI 10.1073/pnas.1211452109.
101. Ravenstein, E. G. The Laws of Migration / E. G. Ravenstein // *Journal of the Statistical Society of London*. – 1885. – Vol. 48, № 2. – P. 167-235.
102. Raymer, J. Applying model migration schedules to represent age-specific migration flows / J. Raymer, A. Rogers // *International Migration in Europe: Data, Models and Estimates* / ed. by J. Raymer, F. Willekens. – Chichester : John Wiley & Sons, 2008. – P. 175-192. – DOI 10.1002/9780470985557.ch8.
103. Renshaw, A. E. A cohort-based extension to the Lee-Carter model for mortality reduction factors / A. E. Renshaw, S. Haberman // *Insurance: Mathematics and Economics*. – 2006. – Vol. 38, № 3. – P. 556-570. – DOI 10.1016/j.insmatheco.2005.12.001.

104. Schoemaker, P. J. H. Scenario Planning: A Tool for Strategic Thinking / P. J. H. Schoemaker // Sloan Management Review. – 1995. – Vol. 36, № 2. – P. 25-40.
105. Smith, S. K. State and Local Population Projections: Methodology and Analysis / S. K. Smith, J. Tayman, D. A. Swanson. – New York : Kluwer Academic Publishers, 2005. – 426 p.
106. Spanish National Statistics Institute. Population Projections of Spain 2022-2072. Methodology. – Madrid, 2022.
107. Vanella, P. A probabilistic cohort-component population projection for Germany / P. Vanella, P. Deschermeier // Journal of Population Research. – 2018. – Vol. 35. – P. 87-118.
108. Willekens, F. J. A multiregional model for regional demographic projection / F. J. Willekens, P. Drewe // Readings in population research methodology / ed. by J. Bogue, E. E. Arriaga et al. – Chicago ; New York : Social Development Center/UNFPA, 1993. – Vol. 5. – P. 17-49-17-57.
109. Wilson, T. Forecast accuracy and uncertainty of Australian Bureau of Statistics state and territory population projections / T. Wilson // International Journal of Population Research. – 2012. – Article ID 419824. – DOI 10.1155/2012/419824.
110. Wilson, T. Methods for small area population forecasts: state-of-the-art and research needs / T. Wilson, I. Grossman, M. Alexander, P. Rees, J. Temple // Population Research and Policy Review. – 2022. – Vol. 41. – P. 865-898. – DOI 10.1007/s11113-021-09671-6.
111. Wilson, T. Model migration schedules incorporating student migration peaks / T. Wilson // Demographic Research. – 2010. – Vol. 23, Article 8. – P. 191-222. – DOI 10.4054/DemRes.2010.23.8.
112. Wilson, T. Recent Developments in Population Projection Methodology: A Review / T. Wilson, P. Rees // Population, Space and Place. – 2005. – Vol. 11, № 5. – P. 337-360. – DOI 10.1002/psp.389.
113. World Cities Report 2022: Envisaging the Future of Cities / UN-Habitat. – Nairobi : United Nations Human Settlements Programme, 2022. – 387 p.

114. World Population Prospects 2022: Methodology of the United Nations population estimates and projections / United Nations, Department of Economic and Social Affairs, Population Division. – New York : United Nations, 2022. – UN DESA/POP/2022/TR/NO. 4. – URL: https://population.un.org/wpp/Publications/Files/WPP2022_Methodology.pdf
115. World Population Prospects 2024: Methodology of the United Nations Population Estimates and Projections / United Nations, Department of Economic and Social Affairs, Population Division. – New York : United Nations, 2024. – 86 p. – URL: https://population.un.org/wpp/assets/Files/WPP2024_Methodology-Report_Final.pdf
116. World Urbanization Prospects: The 2018 Revision / United Nations, Department of Economic and Social Affairs, Population Division. – New York : United Nations, 2019. – ST/ESA/SER.A/420. – URL: <https://population.un.org/wup/assets/WUP2018-Report.pdf>
117. Yu, C. C. Probabilistic county-level population projections / C. C. Yu, H. Ševčíková, A. E. Raftery, S. R. Curran // *Demography*. – 2023. – Vol. 60, № 3. – P. 915-937. – DOI 10.1215/00703370-10772782.
118. Ševčíková, H. Probabilistic projection of subnational life expectancy / H. Ševčíková, A. E. Raftery // *Journal of Official Statistics*. – 2021. – Vol. 37, № 4. – P. 921-942.

ПРИЛОЖЕНИЯ

ПРИЛОЖЕНИЕ 1

Таблица 1 – ARIMA(p,d,q) по критерию AIC и приращение относительно ARIMA(1,0,0)

Период калибровки: 1990-2023				
Блок	Пол	k	Лучшая модель	ΔAIC
Fertility	NA	1	(1,1,0)	-7.459
Fertility	NA	2	(0,1,1)	-1.549
Migration	F	1	(1,1,1)	-15.303
Migration	F	2	(1,1,0)	-8.204
Migration	M	1	(0,1,1)	-12.401
Migration	M	2	(0,1,1)	-9.304
Mortality	F	1	(1,1,0)	-4.188
Mortality	F	2	(1,1,1)	-4.937
Mortality	M	1	(1,1,0)	-8.602
Mortality	M	2	(0,1,1)	-2.443
Период калибровки: 1995-2023				
Блок	Пол	k	Лучшая модель	ΔAIC
Fertility	NA	1	(1,1,0)	-7.519
Fertility	NA	2	(0,1,1)	-2.075
Migration	F	1	(1,1,1)	-14.949
Migration	F	2	(1,1,0)	-7.446
Migration	M	1	(0,1,1)	-12.081
Migration	M	2	(0,1,1)	-9.432
Mortality	F	1	(1,1,0)	-5.351
Mortality	F	2	(1,0,0)	0.000
Mortality	M	1	(1,1,1)	-8.324
Mortality	M	2	(1,0,0)	0.000
Период калибровки: 2000-2023				
Блок	Пол	k	Лучшая модель	ΔAIC
Fertility	NA	1	(1,1,0)	-7.514
Fertility	NA	2	(0,1,1)	-0.173
Migration	F	1	(1,1,0)	-9.573
Migration	F	2	(1,1,0)	-8.319
Migration	M	1	(0,1,1)	-10.387
Migration	M	2	(0,1,1)	-10.069
Mortality	F	1	(1,1,0)	-5.317
Mortality	F	2	(1,0,0)	0.000
Mortality	M	1	(1,1,1)	-9.165
Mortality	M	2	(1,0,0)	0.000
Период калибровки: 2005-2023				
Блок	Пол	k	Лучшая модель	ΔAIC
Fertility	NA	1	(1,1,0)	-7.444

Fertility	NA	2	(0,1,1)	-1.300
Migration	F	1	(0,1,1)	-4.586
Migration	F	2	(1,1,0)	-7.097
Migration	M	1	(0,1,1)	-4.556
Migration	M	2	(0,1,1)	-9.936
Mortality	F	1	(1,1,0)	-4.801
Mortality	F	2	(0,1,1)	-0.247
Mortality	M	1	(1,1,1)	-7.229
Mortality	M	2	(0,1,1)	-0.516

Примечание. ΔAIC рассчитывается как разность между AIC лучшей по критерию AIC модели и AIC базовой $ARIMA(1,0,0)$. Отрицательные значения указывают на улучшение по сравнению с $ARIMA(1,0,0)$. Для каждой компоненты ($k=1,2$) и пола модели оценивались отдельно.

ПРИЛОЖЕНИЕ 2

Таблица 1 – Сравнение доли объяснённой дисперсии (PC1 и PC2) при PCA-разложении возрастных профилей на окнах 1990-2023 и 2000-2023 гг.

Компонента	PC1, % (1990-2023)	PC1, % (2000-2023)	PC2, % (1990-2023)	PC2, % (2000-2023)
Смертность (F)	73,88	77,55	12,79	10,09
Смертность (M)	77,63	81,27	11,71	8,62
Миграция (F)	74,05	53,77	10,36	28,88
Миграция (M)	66,83	49,19	14,77	27,87
Рождаемость (F)	87,95	98,78	10,79	1,01

Таблица 1 – Проверка чувствительности предиктивных интервалов к выбору исторического окна калибровки

Окно калибровки	Число стартов backtesting	Число проверок	Среднее покрытие 90%-го интервала	Среднее покрытие 75%-го интервала	Средняя ширина 90%-го интервала, тыс. чел.	Средняя ширина 75%-го интервала, тыс. чел.	Средний k для 90%-го интервала	Средний k для 75%-го интервала
1990-2023	5	25	0,68	0,6	345	140,6	1,2	1,14
1995-2023	5	25	0,64	0,6	933	263,7	1,23	1,19
2000-2023	5	25	0,48	0,44	126,4	89,4	1,41	1,41
2005-2023	5	25	0,48	0,4	120,9	84,8	1,41	1,37

Примечание. Эмпирическое покрытие показывает долю случаев, в которых фактическая численность населения попадала в соответствующий предиктивный интервал при скользящем backtesting на горизонте до 5 лет. Коэффициент k отражает средний множитель растяжки полуинтервала, необходимый для приближения эмпирического покрытия к номинальному уровню.

Таблица 1 – Программная реализация вероятностной когортно-компонентной модели

```
# ===== ЧАСТЬ 1: загрузка данных, проверка структуры и подготовка матриц
=====
import pandas as pd
import numpy as np
from pathlib import Path

# --- параметры
# XLSX: путь к исходной книге с четырьмя листами данных
# HIST_START/HIST_END: границы исторического периода анализа
# START_YEAR: базовый год формирования стартовой возрастно-половой структуры
населения
# REPRO_AGE_GROUPS: стандартные пятилетние репродуктивные возрастные группы для
расчёта ASFR и TFR
XLSX = Path("Population_Forecasting_St_Petersburg.xlsx")
HIST_START, HIST_END = 1990, 2023
START_YEAR = 2023 # базовый год прогноза
REPRO_AGE_GROUPS = ["15-19", "20-24", "25-29", "30-34", "35-39", "40-44", "45-49"]

# --- вспомогательные функции
# norm_sex: приводит обозначения пола к единому формату {'F', 'M'}.
def norm_sex(s):
    s = str(s).strip().lower()
    return {"f": "F", "female": "F", "ж": "F", "m": "M", "male": "M", "м": "M"}.get(s,
s.upper())

# age_key: преобразует метку возрастной группы в числовой ключ для сортировки.
# Поддерживаются интервалы вида 'a-b' и открытая старшая возрастная группа вида
'85+'.
def age_key(g):
    g = str(g).strip()
    if g.endswith('+'):
        base = int(g[:-1])
        return (base, 999)
    a, b = g.split('-')
    return (int(a), int(b))

# ordered_age_cat: задаёт порядок возрастных групп как упорядоченную
категориальную переменную.
# Это необходимо для корректной индексации и последующего переупорядочивания
возрастных групп.
def ordered_age_cat(series):
    uniq = sorted(series.dropna().astype(str).unique(), key=age_key)
    return pd.Categorical(series.astype(str), categories=uniq, ordered=True)

# wide: преобразует длинную таблицу в матричный формат по годам.
# При sex=True формируются составные столбцы (Sex, Age_Group);
# при sex=False столбцы задаются только возрастными группами.
# Возвращается фрагмент данных за HIST_START..HIST_END.
def wide(df, value, sex=False):
    if sex:
        w = df.pivot_table(index="Year", columns=["Sex", "Age_Group"],
values=value,
aggfunc="mean", observed=False)
```

```

w = w.sort_index(axis=1, level=[0,1], sort_remaining=True)
else:
    w = df.pivot_table(index="Year", columns="Age_Group", values=value,
                        aggfunc="mean", observed=False)
    w = w.sort_index(axis=1)
return w.loc[HIST_START:HIST_END]

# --- чтение исходных листов
# Данные импортируются из Excel, после чего поля нормируются и ограничиваются
историческим периодом.
pop = pd.read_excel(XLSX, sheet_name="Base Population Data")
fert = pd.read_excel(XLSX, sheet_name="Fertility Data")
mort = pd.read_excel(XLSX, sheet_name="Mortality Data")
mig = pd.read_excel(XLSX, sheet_name="Migration Data")

# --- нормализация полей
# Sex приводится к формату {'F','M'}, Year — к целочисленному типу,
# Age_Group — к упорядоченной категориальной переменной.
for df in (pop, fert, mort, mig):
    df["Sex"] = df["Sex"].map(norm_sex)
    df["Year"] = df["Year"].astype(int)
    df["Age_Group"] = ordered_age_cat(df["Age_Group"])

# --- фильтрация исторического периода
# Для всех компонент сохраняются только наблюдения за 1990-2023 гг.
pop = pop.query("@HIST_START <= Year <= @HIST_END").copy()
fert = fert.query("@HIST_START <= Year <= @HIST_END").copy()
mort = mort.query("@HIST_START <= Year <= @HIST_END").copy()
mig = mig.query("@HIST_START <= Year <= @HIST_END").copy()

# --- проверка структуры входных данных
# Выполнение прерывается, если в исходных листах отсутствуют обязательные поля.
req_pop = {"Year", "Age_Group", "Sex", "Population"}
req_fert = {"Year", "Age_Group", "Sex", "Fertility_Rate"} # ASFR, ‰ на 1000
женщин
req_mort = {"Year", "Age_Group", "Sex", "Mortality_Rate"} # ASMR, ‰ на 1000
населения
req_mig = {"Year", "Age_Group", "Sex", "Migration_Inflow", "Migration_Outflow"} #
человек

for name, df, req in [
    ("Base Population Data", pop, req_pop),
    ("Fertility Data", fert, req_fert),
    ("Mortality Data", mort, req_mort),
    ("Migration Data", mig, req_mig),
]:
    miss = req - set(df.columns)
    assert not miss, f"{name}: отсутствуют колонки: {sorted(miss)}"

# --- стартовая численность населения
# Формируется матрица Age_Group × Sex для START_YEAR.
# Она используется как начальное состояние когортно-компонентного прогноза.
pop2023 = pop.query("Year == @START_YEAR").copy()
pop_wide = pop2023.pivot_table(index="Age_Group", columns="Sex",
                                values="Population",
                                aggfunc="sum", observed=False).fillna(0)
pop_wide =
pop_wide.reindex(index=ordered_age_cat(pop2023["Age_Group"]).categories)
pop_wide = pop_wide.rename_axis(index="Age_Group", columns="Sex")

```

```

# --- смертность
# mort_w содержит возрастно-половые коэффициенты смертности в формате Year ×
# (Sex, Age_Group).
mort_w = wide(mort, "Mortality_Rate", sex=True)

# --- рождаемость
# Для расчёта возрастных коэффициентов рождаемости используются данные по
# женщинам.
# fert_w содержит ASFR в формате Year × Age_Group.
fert_f = fert.query("Sex == 'F'").copy()
fert_w = wide(fert_f, "Fertility_Rate", sex=False)

# --- миграция
# Формируются матрицы прибытий, выбытий и чистой миграции по полу и возрасту.
# mig_net_w содержит чистую миграцию в формате Year × (Sex, Age_Group).
mig_in_w = wide(mig, "Migration_Inflow", sex=True)
mig_out_w = wide(mig, "Migration_Outflow", sex=True)
mig_net_w = (mig_in_w - mig_out_w)

# --- проверка репродуктивных возрастных групп
# Для устойчивого расчёта TFR требуется наличие не менее пяти стандартных
# пятилетних групп.
have_repro = [g for g in REPRO_AGE_GROUPS if g in
fert_w.columns.astype(str).tolist()]
assert len(have_repro) >= 5, f"ASFR: найдено недостаточно репродуктивных групп:
{have_repro}"

# --- расчёт TFR
# Функция рассчитывает суммарный коэффициент рождаемости как сумму ASFR × 5 /
# 1000,
# поскольку ASFR заданы в промилле, а возрастные группы являются пятилетними.
def compute_TFR_row(asfr_row, group_labels):
    width_years = 5
    vals = asfr_row.reindex(group_labels).astype(float).fillna(0.0)
    tfr = (vals.sum() * width_years) / 1000.0
    return tfr

# --- TFR по годам
# tfr_series содержит оценки TFR за каждый год исторического периода,
# tfr_2023 фиксирует значение TFR в базовом году прогноза.
tfr_series = fert_w.apply(lambda r: compute_TFR_row(r, have_repro), axis=1)
tfr_2023 = float(tfr_series.loc[START_YEAR])

# --- агрегированные показатели за базовый год
# Рассчитываются общая численность населения, численность по полу
# и чистая миграция по полу и в целом.
total_pop_2023 = float(pop2023["Population"].sum())
total_pop_2023_F = float(pop2023.query("Sex=='F'")["Population"].sum())
total_pop_2023_M = float(pop2023.query("Sex=='M'")["Population"].sum())

net_mig_2023_F = float(mig_net_w.xs("F", axis=1, level=0).loc[START_YEAR].sum())
net_mig_2023_M = float(mig_net_w.xs("M", axis=1, level=0).loc[START_YEAR].sum())
net_mig_2023 = net_mig_2023_F + net_mig_2023_M

# --- проверка масштаба коэффициентов смертности
# Рассчитываются средние уровни возрастных коэффициентов смертности по полу
# для базового года прогноза.

```

```

mort_level_2023_F = float(mort_w.xs("F", axis=1,
level=0).loc[START_YEAR].mean())
mort_level_2023_M = float(mort_w.xs("M", axis=1,
level=0).loc[START_YEAR].mean())

# --- контрольный вывод
# Диагностическая печать используется для проверки загруженных данных и
сформированных матриц.
print("=== Диагностика. Часть 1 ===")
print(f"Годы в данных: {HIST_START}-{HIST_END}, базовый год прогноза:
{START_YEAR}")
print(f"Возрастные группы (пример): {list(pop_wide.index[:6])} ...
{list(pop_wide.index[-3:])}")
print(f"Население 2023 всего: {total_pop_2023:,.0f} (F: {total_pop_2023_F:,.0f};
M: {total_pop_2023_M:,.0f})")
print(f"Чистая миграция 2023: {net_mig_2023:,.0f} (F: {net_mig_2023_F:,.0f}; M:
{net_mig_2023_M:,.0f})")
print(f"TFR(2023) по ASFR (%): {tfr_2023:.3f}")
print(f"Среднее по возрастным коэффициентам смертности (%), 2023:
F={mort_level_2023_F:.3f}, M={mort_level_2023_M:.3f}")
print(f"ASFR группы в работе: {have_repro}")
print("Матрицы подготовлены: mort_w (Year × Sex × Age), fert_w (Year × Age),
mig_net_w (Year × Sex × Age), pop_wide (Age × Sex)")
# =====

```

```

# ===== ЧАСТЬ 1: загрузка данных, проверка структуры и подготовка матриц
=====
import pandas as pd
import numpy as np
from pathlib import Path

# --- параметры
# XLSX: путь к исходной книге с четырьмя листами данных
# HIST_START/HIST_END: границы исторического периода анализа
# START_YEAR: базовый год формирования стартовой возрастно-половой структуры
населения
# REPRO_AGE_GROUPS: стандартные пятилетние репродуктивные возрастные группы для
расчёта ASFR и TFR
XLSX = Path("Population_Forecasting_St_Petersburg.xlsx")
HIST_START, HIST_END = 1990, 2023
START_YEAR = 2023 # базовый год прогноза
REPRO_AGE_GROUPS = ["15-19", "20-24", "25-29", "30-34", "35-39", "40-44", "45-49"]

# --- вспомогательные функции
# norm_sex: приводит обозначения пола к единому формату {'F', 'M'}.
def norm_sex(s):
    s = str(s).strip().lower()
    return {"f": "F", "female": "F", "ж": "F", "m": "M", "male": "M", "м": "M"}.get(s,
s.upper())

# age_key: преобразует метку возрастной группы в числовой ключ для сортировки.
# Поддерживаются интервалы вида 'a-b' и открытая старшая возрастная группа вида
'85+'.
def age_key(g):
    g = str(g).strip()
    if g.endswith('+'):
        base = int(g[:-1])
        return (base, 999)
    a, b = g.split('-')

```

```

return (int(a), int(b))

# ordered_age_cat: задаёт порядок возрастных групп как упорядоченную
# категориальную переменную.
# Это необходимо для корректной индексации и последующего переупорядочивания
# возрастных групп.
def ordered_age_cat(series):
    uniq = sorted(series.dropna().astype(str).unique(), key=age_key)
    return pd.Categorical(series.astype(str), categories=uniq, ordered=True)

# wide: преобразует длинную таблицу в матричный формат по годам.
# При sex=True формируются составные столбцы (Sex, Age_Group);
# при sex=False столбцы задаются только возрастными группами.
# Возвращается фрагмент данных за HIST_START..HIST_END.
def wide(df, value, sex=False):
    if sex:
        w = df.pivot_table(index="Year", columns=["Sex", "Age_Group"],
        values=value,
                                aggfunc="mean", observed=False)
        w = w.sort_index(axis=1, level=[0,1], sort_remaining=True)
    else:
        w = df.pivot_table(index="Year", columns="Age_Group", values=value,
                                aggfunc="mean", observed=False)
        w = w.sort_index(axis=1)
    return w.loc[HIST_START:HIST_END]

# --- чтение исходных листов
# Данные импортируются из Excel, после чего поля нормируются и ограничиваются
# историческим периодом.
pop = pd.read_excel(XLSX, sheet_name="Base Population Data")
fert = pd.read_excel(XLSX, sheet_name="Fertility Data")
mort = pd.read_excel(XLSX, sheet_name="Mortality Data")
mig = pd.read_excel(XLSX, sheet_name="Migration Data")

# --- нормализация полей
# Sex приводится к формату {'F','M'}, Year — к целочисленному типу,
# Age_Group — к упорядоченной категориальной переменной.
for df in (pop, fert, mort, mig):
    df["Sex"] = df["Sex"].map(norm_sex)
    df["Year"] = df["Year"].astype(int)
    df["Age_Group"] = ordered_age_cat(df["Age_Group"])

# --- фильтрация исторического периода
# Для всех компонент сохраняются только наблюдения за 1990–2023 гг.
pop = pop.query("@HIST_START <= Year <= @HIST_END").copy()
fert = fert.query("@HIST_START <= Year <= @HIST_END").copy()
mort = mort.query("@HIST_START <= Year <= @HIST_END").copy()
mig = mig.query("@HIST_START <= Year <= @HIST_END").copy()

# --- проверка структуры входных данных
# Выполнение прерывается, если в исходных листах отсутствуют обязательные поля.
req_pop = {"Year", "Age_Group", "Sex", "Population"}
req_fert = {"Year", "Age_Group", "Sex", "Fertility_Rate"} # ASFR, ‰ на 1000
# женщин
req_mort = {"Year", "Age_Group", "Sex", "Mortality_Rate"} # ASMR, ‰ на 1000
# населения
req_mig = {"Year", "Age_Group", "Sex", "Migration_Inflow", "Migration_Outflow"} #
# человек

```

```

for name, df, req in [
    ("Base Population Data", pop, req_pop),
    ("Fertility Data",      fert, req_fert),
    ("Mortality Data",      mort, req_mort),
    ("Migration Data",      mig, req_mig),
]:
    miss = req - set(df.columns)
    assert not miss, f"{name}: отсутствуют колонки: {sorted(miss)}"

# --- стартовая численность населения
# Формируется матрица Age_Group × Sex для START_YEAR.
# Она используется как начальное состояние когортно-компонентного прогноза.
pop2023 = pop.query("Year == @START_YEAR").copy()
pop_wide = pop2023.pivot_table(index="Age_Group", columns="Sex",
                               values="Population",
                               aggfunc="sum", observed=False).fillna(0)
pop_wide =
pop_wide.reindex(index=ordered_age_cat(pop2023["Age_Group"]).categories)
pop_wide = pop_wide.rename_axis(index="Age_Group", columns="Sex")

# --- смертность
# mort_w содержит возрастно-половые коэффициенты смертности в формате Year ×
# (Sex, Age_Group).
mort_w = wide(mort, "Mortality_Rate", sex=True)

# --- рождаемость
# Для расчёта возрастных коэффициентов рождаемости используются данные по
# женщинам.
# fert_w содержит ASFR в формате Year × Age_Group.
fert_f = fert.query("Sex == 'F']").copy()
fert_w = wide(fert_f, "Fertility_Rate", sex=False)

# --- миграция
# Формируются матрицы прибытий, выбытий и чистой миграции по полу и возрасту.
# mig_net_w содержит чистую миграцию в формате Year × (Sex, Age_Group).
mig_in_w = wide(mig, "Migration_Inflow", sex=True)
mig_out_w = wide(mig, "Migration_Outflow", sex=True)
mig_net_w = (mig_in_w - mig_out_w)

# --- проверка репродуктивных возрастных групп
# Для устойчивого расчёта TFR требуется наличие не менее пяти стандартных
# пятилетних групп.
have_repro = [g for g in REPRO_AGE_GROUPS if g in
fert_w.columns.astype(str).tolist()]
assert len(have_repro) >= 5, f"ASFR: найдено недостаточно репродуктивных групп:
{have_repro}"

# --- расчёт TFR
# Функция рассчитывает суммарный коэффициент рождаемости как сумму ASFR × 5 /
1000,
# поскольку ASFR заданы в промилле, а возрастные группы являются пятилетними.
def compute_TFR_row(asfr_row, group_labels):
    width_years = 5
    vals = asfr_row.reindex(group_labels).astype(float).fillna(0.0)
    tfr = (vals.sum() * width_years) / 1000.0
    return tfr

# --- TFR по годам
# tfr_series содержит оценки TFR за каждый год исторического периода,

```

```

# tfr_2023 фиксирует значение TFR в базовом году прогноза.
tfr_series = fert_w.apply(lambda r: compute_TFR_row(r, have_repro), axis=1)
tfr_2023 = float(tfr_series.loc[START_YEAR])

# --- агрегированные показатели за базовый год
# Рассчитываются общая численность населения, численность по полу
# и чистая миграция по полу и в целом.
total_pop_2023 = float(pop2023["Population"].sum())
total_pop_2023_F = float(pop2023.query("Sex=='F'")["Population"].sum())
total_pop_2023_M = float(pop2023.query("Sex=='M'")["Population"].sum())

net_mig_2023_F = float(mig_net_w.xs("F", axis=1, level=0).loc[START_YEAR].sum())
net_mig_2023_M = float(mig_net_w.xs("M", axis=1, level=0).loc[START_YEAR].sum())
net_mig_2023 = net_mig_2023_F + net_mig_2023_M

# --- проверка масштаба коэффициентов смертности
# Рассчитываются средние уровни возрастных коэффициентов смертности по полу
# для базового года прогноза.
mort_level_2023_F = float(mort_w.xs("F", axis=1, level=0).loc[START_YEAR].mean())
mort_level_2023_M = float(mort_w.xs("M", axis=1, level=0).loc[START_YEAR].mean())

# --- контрольный вывод
# Диагностическая печать используется для проверки загруженных данных и
# сформированных матриц.
print("=== Диагностика. Часть 1 ===")
print(f"Годы в данных: {HIST_START}-{HIST_END}, базовый год прогноза: {START_YEAR}")
print(f"Возрастные группы (пример): {list(pop_wide.index[:6])} ... {list(pop_wide.index[-3:])}")
print(f"Население 2023 всего: {total_pop_2023:,.0f} (F: {total_pop_2023_F:,.0f}; M: {total_pop_2023_M:,.0f})")
print(f"Чистая миграция 2023: {net_mig_2023:,.0f} (F: {net_mig_2023_F:,.0f}; M: {net_mig_2023_M:,.0f})")
print(f"TFR(2023) по ASFR (%): {tfr_2023:.3f}")
print(f"Среднее по возрастным коэффициентам смертности (%), 2023: F={mort_level_2023_F:.3f}, M={mort_level_2023_M:.3f}")
print(f"ASFR группы в работе: {have_repro}")
print(f"Матрицы подготовлены: mort_w (Year × Sex × Age), fert_w (Year × Age), mig_net_w (Year × Sex × Age), pop_wide (Age × Sex)")
# =====

```

```

# ===== ЧАСТЬ 3: Монте-Карло-симуляция когортно-компонентного прогноза =====
import numpy as np
import pandas as pd

# Ожидается, что уже сформированы pop_wide, fert_w, mort_w, mig_net_w,
# mort_models, fert_model, mig_models, а также mig_in_w.
# Если mig_in_w отсутствует, он формируется ниже.

YEARS = list(range(2024, 2046)) # прогнозный период 2024-2045 гг.
SEXES = ["F", "M"]
AGE_LABELS = [str(a) for a in pop_wide.index] # возрастные группы в исходной базе
ASFR_LABELS = [str(a) for a in fert_w.columns] # возрастные группы ASFR
REPRO AGE GROUPS = ["15-19", "20-24", "25-29", "30-34", "35-39", "40-44", "45-49"]

```

```

# ---- основные параметры модели -----
SRB = 1.06                                # соотношение полов при рождении, М/Ф
MIG_GROUP_CAP = 8000.0                    # ограничение миграции на одну возрастную
группу, чел./год
MIG_NET_BUMP = 8_000                       # добавка к чистой миграции в
трудоспособных возрастах, чел./год
MIG_BUMP_SEX_SPLIT = (0.48, 0.52)         # распределение добавки по полу, F/M
OLD_AGE_MORT_MULT = 1.05                  # корректировка смертности в группе 85+

# ---- возрастные ограничения коэффициентов смертности, промилле -----
AGE_LO = [int(a.split("-")[0].replace("+", "")) for a in AGE_LABELS]
CAP_BY_AGE = np.array([
    35.0 if a < 50 else
    55.0 if a < 65 else
    95.0 if a < 75 else
    160.0 if a < 80 else
    260.0 if a < 85 else
    340.0
    for a in AGE_LO
], dtype=float)

# ---- распределение добавки к чистой миграции по возрастам 20-44 года -----
# Максимальный вес задаётся для групп 25-34 года.
bump_weights = pd.Series(0.0, index=AGE_LABELS, dtype=float)
for ag in AGE_LABELS:
    a0 = int(ag.split("-")[0].replace("+", ""))
    if 20 <= a0 <= 44:
        if a0 < 25: w = (a0 - 20) / 5
        elif a0 <= 34: w = 1.0
        else: w = max(0.0, 1 - (a0 - 35)/10)
        bump_weights[ag] = w
if bump_weights.sum() == 0: bump_weights += 1.0
bump_weights /= bump_weights.sum()

# ---- эмпирические возрастные веса миграции по полу, 2014-2023 гг. -----
# Используется экспоненциально взвешенное среднее с параметром lambda = 0.8.
# Более поздние годы получают больший вес.
try:
    mig_in_w
except NameError:
    mig_in_w = (mig.pivot_table(index="Year", columns=["Sex", "Age_Group"],
                                values="Migration_Inflow", aggfunc="sum",
                                observed=False)
                .sort_index(axis=1))

def _normalize_floor(s, floor=1e-9):
    s = s.astype(float).fillna(0.0).clip(lower=0.0)
    if s.sum() <= 0: s += floor
    s[s < floor] = floor
    return s / s.sum()

def _age_shares_per_year(sub, sex):
    mat = sub.xs(sex, axis=1, level=0)
    sums = mat.sum(axis=1).replace(0, np.nan)
    return mat.div(sums, axis=0)

# Окно расчёта и веса по годам.
years_all = sorted(mig_in_w.index.tolist())

```

```

y0, y1 = 2014, 2023
years_window = [y for y in years_all if y0 <= y <= y1]
sub = mig_in_w.loc[years_window]

sharesF = _age_shares_per_year(sub, "F")
sharesM = _age_shares_per_year(sub, "M")

lam = 0.8
yrs = sharesF.index.to_numpy()
tmax = yrs.max()
w_year = np.array([lam**(tmax - t) for t in yrs], dtype=float); w_year /=
w_year.sum()

def _ewa(shares, weights):
    arr = shares.to_numpy(dtype=float)
    wcol = weights.reshape(-1,1)
    num = np.nansum(arr * wcol, axis=0)
    den = np.nansum((~np.isnan(arr)) * wcol, axis=0)
    out = np.divide(num, den, out=np.zeros_like(num), where=den>0)
    return pd.Series(out, index=shares.columns).fillna(0.0)

wF_emp = _normalize_floor(_ewa(sharesF, w_year))
wM_emp = _normalize_floor(_ewa(sharesM, w_year))

# Приведение порядка возрастных групп к AGE_LABELS.
wF = wF_emp.reindex(AGE_LABELS).fillna(0.0).to_numpy(float)
wM = wM_emp.reindex(AGE_LABELS).fillna(0.0).to_numpy(float)

# ---- функции годового перехода, дожития, рождаемости и миграции -----
def surv_prob_exp_by_age(mort_rate_promille_vec):
    """Годовая вероятность дожития по возрастам:  $p = \exp(-m/1000)$ .
    Для группы 85+ применяется корректирующий множитель смертности."""
    m = np.asarray(mort_rate_promille_vec, float)
    m = m * np.where(np.array(AGE_LO) >= 85, OLD_AGE_MORT_MULT, 1.0)
    mx = np.clip(m, 0.0, CAP_BY_AGE) / 1000.0
    return np.exp(-mx)

def build_annual_transfer(n_age):
    """Матрица годового перехода между пятилетними возрастными группами:
    4/5 численности остаётся в группе, 1/5 переходит в следующую группу.
    Старшая возрастная группа является открытой."""
    Tm = np.zeros((n_age, n_age))
    for i in range(n_age):
        stay = 1.0 if i == n_age - 1 else 4.0/5.0
        Tm[i, i] += stay
        if i < n_age - 1:
            Tm[i+1, i] += 1.0/5.0
    return Tm

A = build_annual_transfer(len(AGE_LABELS))
AGE_TRANSFER = A

def births_from_asfr(pop_f_by_age, asfr_promille_by_age):
    """Число рождений за год рассчитывается как сумма произведений
    ASFR и численности женщин соответствующих возрастных групп, делённая на
    1000."""
    popF = pd.Series(pop_f_by_age, index=AGE_LABELS, dtype=float)
    asfr = pd.Series(asfr_promille_by_age, index=ASFR_LABELS,
dtype=float).fillna(0.0)

```

```

births = (asfr.reindex(REPRO_AGE_GROUPS).fillna(0.0) *
          popF.reindex(REPRO_AGE_GROUPS).fillna(0.0)).sum() / 1000.0
return max(births, 0.0)

def shape_migration_sex(vec, sex, cap=MIG_GROUP_CAP):
    """Возрастное распределение миграции с использованием эмпирических весов по
    полу.

    После взвешивания сохраняется общий объём миграции и применяется
    ограничение по группе."""
    v = np.asarray(vec, float)
    W = wF if sex == "F" else wM
    v_w = v * W
    s, sw = v.sum(), v_w.sum()
    if abs(sw) < 1e-12:
        return np.zeros_like(v)
    v_adj = v_w * (s / sw)
    return np.clip(v_adj, -cap, cap)

# --- проверка согласованности возрастных порядков -----
for s in SEXES:
    assert [str(c) for c in mort_w.xs(s, axis=1, level=0).columns] == AGE_LABELS
    assert [str(c) for c in mig_net_w.xs(s, axis=1, level=0).columns] ==
AGE_LABELS
assert [str(c) for c in fert_w.columns] == ASFR_LABELS

# --- Монте-Карло-симуляции -----
N = N_SIMS
T = len(YEARS); A_LEN = len(AGE_LABELS)
p_male = SRB / (1 + SRB)
p_fem = 1 - p_male

pop_sim_F = np.zeros((N, T, A_LEN))
pop_sim_M = np.zeros((N, T, A_LEN))

# Стартовые возрастно-половые векторы населения на конец 2023 г.
base_F = pop_wide["F"].astype(float).values.copy()
base_M = pop_wide["M"].astype(float).values.copy()

for d in range(N):
    popF = base_F.copy()
    popM = base_M.copy()

    for t in range(T):
        # Стохастические возрастные профили демографических компонент на год t.
        mortF = mort_models["F"]["sims"][d, t, :]
        mortM = mort_models["M"]["sims"][d, t, :]
        migF = mig_models["F"]["sims"][d, t, :]
        migM = mig_models["M"]["sims"][d, t, :]
        asfr = fert_model["sims"][d, t, :]

        # Смертность, дожитие и годовой переход между возрастными группами.
        survF = surv_prob_exp_by_age(mortF)
        survM = surv_prob_exp_by_age(mortM)
        popF = A @ (popF * survF)
        popM = A @ (popM * survM)

        # Миграция с учётом эмпирической возрастной структуры и ограничения по
        группе.
        popF += shape_migration_sex(migF, "F", cap=MIG_GROUP_CAP)

```

```

popM += shape_migration_sex(migM, "M", cap=MIG_GROUP_CAP)

# Уровневая добавка к чистой миграции в трудоспособных возрастах.
bumpF = (MIG_NET_BUMP * MIG_BUMP_SEX_SPLIT[0]) * bump_weights.values
bumpM = (MIG_NET_BUMP * MIG_BUMP_SEX_SPLIT[1]) * bump_weights.values
popF += bumpF
popM += bumpM

# Рождения и распределение новорождённых по полу.
B = births_from_asfr(popF, asfr)
popF[0] += B * p_fem
popM[0] += B * p_male

pop_sim_F[d, t, :] = popF
pop_sim_M[d, t, :] = popM

# --- агрегирование результатов и диагностический вывод -----
totF = pop_sim_F.sum(axis=2)
totM = pop_sim_M.sum(axis=2)
totAll = totF + totM

median_pop = np.median(totAll, axis=0)
p10 = np.percentile(totAll, 10, axis=0)
p90 = np.percentile(totAll, 90, axis=0)

t_2045 = YEARS.index(2045)
share85 = (np.median(pop_sim_F[:, t_2045, -1] + pop_sim_M[:, t_2045, -1],
axis=0)) / median_pop[-1]

forecast_df = pd.DataFrame({"Year": YEARS, "Median": median_pop, "P10": p10,
"P90": p90})

print("=== Диагностика. Часть 3 ===")
print(forecast_df.tail())
print(f"Итог 2045: медиана={median_pop[-1]:.0f}, 80% интервал=({p10[-1]:.0f}-
{p90[-1]:.0f})")
print(f"Доля 85+ в 2045 (медиана): {share85*100:.1f}%")
#
=====#

```

```

# === ЧАСТЬ А: Backtest-калибровка предиктивных интервалов и применение к
прогнозу ===
# Требуются результаты предыдущих частей: pop, mort, fert, mig, AGE_LABELS,
# REPRO_AGE_GROUPS, YEARS, матрица A, функции surv_prob_exp_by_age,
# births_from_asfr, shape_migration_sex, а также SRB, MIG_GROUP_CAP и totAll.

import numpy as np
import pandas as pd
from sklearn.decomposition import PCA
from statsmodels.tsa.arima.model import ARIMA
import warnings

# --- подавление технических предупреждений pandas и statsmodels ---
warnings.filterwarnings("ignore", category=FutureWarning)
warnings.filterwarnings("ignore", category=UserWarning)
warnings.filterwarnings("ignore", category=RuntimeWarning)

# =====
# Параметры калибровки

```

```

# =====
BT_ORIGINS = tuple(range(1991, 2020, 2))      # стартовые годы ретроспективной
проверки
BT_HORIZON = 5                                # горизонт ретроспективной проверки,
лет
BT_DRAWS = 1000                               # число симуляций для backtest
MIG_TRANSFORM = "asinh"                       # преобразование для знакопеременных
рядов чистой миграции
MIG_SCALE = 6000.0                           # параметр масштабирования для
asinh-преобразования
RESID_INFLATE_TAU = 1.4                      # множитель стандартного отклонения
остатков ARIMA
K_FLOOR, K_CAP = 1.0, 3.5                   # нижняя и верхняя границы
множителей растяжки
TARGET_COVERAGE = {"90": 0.90, "75": 0.75} # целевые уровни покрытия

# =====
# Вспомогательные функции
# =====
def wide_until(df, value, sex=False, end_year=None):
    """Формирование wide-матрицы с обрезкой истории по end_year включительно."""
    d = df.copy()
    if end_year is not None:
        d = d.loc[d["Year"] <= int(end_year)].copy()
    if sex:
        w = d.pivot_table(index="Year", columns=["Sex", "Age_Group"],
values=value,
                                aggfunc="mean", observed=False).sort_index(axis=1,
level=[0,1], sort_remaining=True)
    else:
        w = d.pivot_table(index="Year", columns="Age_Group", values=value,
                                aggfunc="mean", observed=False).sort_index(axis=1)
    return w

def pca_arima_any(wide_df, n_comp=2, arima_order=(1,0,0), steps=5, draws=1000,
transform="log", scale=2000.0, tau=1.0):
    """
    Универсальная функция PCA+ARIMA для рядов с различными преобразованиями.
    Возвращает стохастические траектории и среднюю прогнозную траекторию
    в исходных единицах измерения.

    transform="log" используется для положительных рядов смертности и
    рождаемости.
    transform="asinh" используется для знакопеременных рядов чистой миграции.
    transform="none" оставляет значения без преобразования.
    tau задаёт множитель стандартного отклонения остатков.
    """
    V = wide_df.values.astype(float)
    if transform == "log":
        X = np.log(np.maximum(V, 1e-6)); inv = np.exp
    elif transform == "asinh":
        X = np.arcsinh(V / scale);          inv = lambda Z: scale * np.sinh(Z)
    else:
        X = V;                              inv = lambda Z: Z

    pca = PCA(n_components=min(n_comp, X.shape[1])).fit(X)
    S = pca.transform(X)
    steps = int(steps)

```

```

S_fore = np.zeros((steps, pca.n_components_))
S_sims = np.zeros((draws, steps, pca.n_components_))

for i in range(pca.n_components_):
    try:
        res = ARIMA(S[:, i], order=arima_order, trend="t").fit()
        base = res.forecast(steps=steps)
        sigma = float(np.std(res.resid)) or 0.1
    except Exception:
        base = np.repeat(S[-1, i], steps)
        sigma = 0.1

    S_fore[:, i] = base
    S_sims[:, :, i] = base[None, :] + np.random.normal(0.0, sigma*tau,
size=(draws, steps))

recon = lambda s: inv(s @ pca.components_ + pca.mean_)
mean_fore = recon(S_fore)
sims = np.array([recon(S_sims[d]) for d in range(draws)])

return {"sims": sims, "mean_fore": mean_fore}

# =====
# Ретроспективная проверка покрытия
# =====
def backtest_coverage(origins=BT_ORIGINS, horizon=BT_HORIZON, draws=BT_DRAWS,
                    mig_transform=MIG_TRANSFORM, mig_scale=MIG_SCALE,
tau=RESID_INFLATE_TAU):
    """
    Для каждого стартового года модели обучаются на истории до origin
    включительно,
    после чего строится прогноз на horizon лет. Возвращается таблица с
    фактическими
    значениями, медианой прогноза, интервальными границами и признаками
    попадания
    факта в 90% и 75% интервалы.

    В ретроспективной проверке не используется MIG_NET_BUMP, чтобы оценить
    собственную неопределённость стохастической модели.
    """
    results = []
    p_male = SRB / (1 + SRB) # доля мальчиков при рождении

    for origin in origins:
        # Формирование wide-матриц компонент до стартового года backtest.
        mortF_w = wide_until(mort, "Mortality_Rate", sex=True,
end_year=origin).xs("F", axis=1, level=0)
        mortM_w = wide_until(mort, "Mortality_Rate", sex=True,
end_year=origin).xs("M", axis=1, level=0)
        fertF_w = wide_until(fert.query("Sex=='F'"), "Fertility_Rate",
sex=False, end_year=origin)
        migF_w = (wide_until(mig, "Migration_Inflow", sex=True,
end_year=origin).xs("F", axis=1, level=0)
- wide_until(mig, "Migration_Outflow", sex=True,
end_year=origin).xs("F", axis=1, level=0))
        migM_w = (wide_until(mig, "Migration_Inflow", sex=True,
end_year=origin).xs("M", axis=1, level=0)
- wide_until(mig, "Migration_Outflow", sex=True,
end_year=origin).xs("M", axis=1, level=0))

```

```

# Стартовая возрастно-половая структура населения на год origin.
base_pop = pop.query("Year == @origin").copy()
base_F = (base_pop.query("Sex=='F'")
          .pivot_table(index="Age_Group", values="Population",
aggfunc="sum")
          .reindex(AGE_LABELS).fillna(0.0).values.ravel())
base_M = (base_pop.query("Sex=='M'")
          .pivot_table(index="Age_Group", values="Population",
aggfunc="sum")
          .reindex(AGE_LABELS).fillna(0.0).values.ravel())

# Моделирование компонент.
# Для смертности и рождаемости используется log-преобразование,
# для чистой миграции используется asinh-преобразование.
mortF_m = pca_arima_any(mortF_w, steps=horizon, draws=draws,
transform="log", tau=tau)
mortM_m = pca_arima_any(mortM_w, steps=horizon, draws=draws,
transform="log", tau=tau)
fert_m = pca_arima_any(fertF_w, steps=horizon, draws=draws,
transform="log", tau=tau)
migF_m = pca_arima_any(migF_w, steps=horizon, draws=draws,
transform=mig_transform, scale=mig_scale, tau=tau)
migM_m = pca_arima_any(migM_w, steps=horizon, draws=draws,
transform=mig_transform, scale=mig_scale, tau=tau)

# Монте-Карло-симуляция: смертность, годовой переход, миграция и
рождения.
# Уровневая добавка к чистой миграции в backtest не применяется.
N, T = draws, horizon
simF = np.zeros((N, T, len(AGE_LABELS))); simM = np.zeros_like(simF)

for d in range(N):
    popF, popM = base_F.copy(), base_M.copy()

    for t in range(T):
        mortF_t = mortF_m["sims"][d, t, :]; mortM_t = mortM_m["sims"][d,
t, :]
        migF_t = migF_m["sims"][d, t, :]; migM_t = migM_m["sims"][d,
t, :]
        asfr_t = fert_m["sims"][d, t, :]

        survF = surv_prob_exp_by_age(mortF_t); survM =
surv_prob_exp_by_age(mortM_t)
        popF = A @ (popF * survF); popM = A @ (popM * survM)

        popF += shape_migration_sex(migF_t, "F", cap=MIG_GROUP_CAP)
        popM += shape_migration_sex(migM_t, "M", cap=MIG_GROUP_CAP)

        B = births_from_asfr(popF, asfr_t)
        popF[0] += B * (1 - p_male); popM[0] += B * p_male

        simF[d, t, :] = popF; simM[d, t, :] = popM

tot = simF.sum(axis=2) + simM.sum(axis=2)
years_chk = list(range(origin+1, origin+1+T))

fact = (pop.query("Year in @years_chk")
        .groupby("Year")["Population"].sum())

```

```

        .reindex(years_chk).values)

    rows = []
    for k, y in enumerate(years_chk):
        q5, q95 = np.percentile(tot[:, k], [5, 95])
        q12, q88 = np.percentile(tot[:, k], [12.5, 87.5])
        rows.append((origin, y, fact[k], np.median(tot[:, k]), q5, q95, q12,
q88))

    df_part = pd.DataFrame(rows,
columns=["Origin", "Year", "Fact_Total", "Median", "P5", "P95", "P12_5", "P87_5"])
    df_part["Covered90"] = (df_part["P5"] <= df_part["Fact_Total"]) &
(df_part["Fact_Total"] <= df_part["P95"])
    df_part["Covered75"] = (df_part["P12_5"] <= df_part["Fact_Total"]) &
(df_part["Fact_Total"] <= df_part["P87_5"])

    results.append(df_part)

res_df = pd.concat(results, ignore_index=True)

# Сводка эмпирического покрытия по горизонту прогноза.
tmp = res_df.copy(); tmp["H"] = tmp["Year"] - tmp["Origin"]
cov90_byH = tmp.groupby("H")["Covered90"].mean()
cov75_byH = tmp.groupby("H")["Covered75"].mean()

print("\n=== Резюме калибровки: эмпирическое покрытие ===")
print("90% интервал по горизонтам:\n", cov90_byH)
print("75% интервал по горизонтам:\n", cov75_byH)
print("Среднее покрытие 90%:", cov90_byH.mean().round(3), "| 75%:",
cov75_byH.mean().round(3))

    return res_df

# =====
# Расчёт множителей калибровки и их применение
# =====
def calib_factors_by_horizon(backtest_df, target_cov=TARGET_COVERAGE,
floor=K_FLOOR, cap=K_CAP):
    """Возвращает словарь Series: {'90': k90(H), '75': k75(H)} по горизонту
H."""
    tmp = backtest_df.copy(); tmp["H"] = tmp["Year"] - tmp["Origin"]
    out = {}

    for lab, tgt in target_cov.items():
        col = "Covered90" if lab=="90" else "Covered75"
        emp_byH = tmp.groupby("H")[col].mean().replace(0, 1e-
6).clip(upper=0.999)
        k_byH = np.sqrt(tgt / emp_byH).clip(lower=floor, upper=cap)
        out[lab] = k_byH

    return out

def apply_calibration_by_horizon(total_df, k90_byH, k75_byH, years=YEARS):
    """Растяжка интервалов вокруг медианы с использованием k(H) для каждого
горизонта."""
    H = np.array([y - years[0] + 1 for y in years], dtype=int)

    def stretch_by_kH(lo, hi, k_series):
        k_vec = np.array([k_series.get(h, k_series.iloc[-1]) for h in H])

```

```

        mid = (hi + lo) / 2.0
        half = (hi - lo) / 2.0
        return mid - half * k_vec, mid + half * k_vec

    out = total_df.copy()
    out["P5"], out["P95"] = stretch_by_kH(out["P5"].values,
out["P95"].values, k90_byH)
    out["P12_5"], out["P87_5"] = stretch_by_kH(out["P12_5"].values,
out["P87_5"].values, k75_byH)

    return out

# =====
# Запуск калибровки
# =====
backtest_df = backtest_coverage(
    origins=BT_ORIGINS, horizon=BT_HORIZON, draws=BT_DRAWS,
    mig_transform=MIG_TRANSFORM, mig_scale=MIG_SCALE, tau=RESID_INFLATE_TAU
)

k_byH = calib_factors_by_horizon(backtest_df, target_cov=TARGET_COVERAGE,
floor=K_FLOOR, cap=K_CAP)

print("\nМножители калибровки по горизонту:")
print("k90 (H): \n", k_byH["90"])
print("k75 (H): \n", k_byH["75"])

# Формирование таблицы итогового прогноза по массиву стохастических реализаций
totAll.
p5 = np.percentile(totAll, 5, axis=0)
p95 = np.percentile(totAll, 95, axis=0)
p12 = np.percentile(totAll, 12.5, axis=0)
p87 = np.percentile(totAll, 87.5, axis=0)
median_pop = np.median(totAll, axis=0)

total_df = pd.DataFrame({
    "Year": YEARS,
    "Median": median_pop,
    "P5": p5,
    "P95": p95,
    "P12_5": p12,
    "P87_5": p87
})

# Применение калибровки по горизонту прогноза.
total_df_CAL = apply_calibration_by_horizon(total_df, k_byH["90"], k_byH["75"],
years=YEARS)

print("\n=== Итог после калибровки по горизонту прогноза ===")
print(total_df_CAL.tail())
print(
    f"2045: медиана =
{total_df_CAL.loc[total_df_CAL['Year']==2045, 'Median'].values[0]:,.0f}, "
    f"90% интервал =
({total_df_CAL.loc[total_df_CAL['Year']==2045, 'P5'].values[0]:,.0f} - "
    f"{total_df_CAL.loc[total_df_CAL['Year']==2045, 'P95'].values[0]:,.0f}), "
    f"75% интервал =
({total_df_CAL.loc[total_df_CAL['Year']==2045, 'P12_5'].values[0]:,.0f} - "
    f"{total_df_CAL.loc[total_df_CAL['Year']==2045, 'P87_5'].values[0]:,.0f}) "

```

```
)

backtest_df.to_excel("backtest_coverage_intervals.xlsx", index=False)
total_df.to_excel("population_forecast_intervals_raw.xlsx", index=False)
total_df_CAL.to_excel("population_forecast_intervals_calibrated.xlsx",
index=False)

print("Сохранены файлы:")
print("backtest_coverage_intervals.xlsx")
print("population_forecast_intervals_raw.xlsx")
print("population_forecast_intervals_calibrated.xlsx")
```

```
# =====
# ЧАСТЬ В: выгрузка и визуализация результатов прогноза
# =====

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

# =====
# 1. Общая численность населения
# Используется таблица total_df_CAL, сформированная после калибровки интервалов.
# =====

total_df_CAL.to_excel("forecast_total_with_PI.xlsx", index=False)

print("Файл сохранён: forecast_total_with_PI.xlsx")
display(total_df_CAL.head())

# =====
# 2. Возрастно-половая структура населения
# Используются массивы pop_sim_F и pop_sim_M, полученные в Монте-Карло-
# симуляции.
# =====

popF_median = np.median(pop_sim_F, axis=0)
popM_median = np.median(pop_sim_M, axis=0)

popF_p10 = np.percentile(pop_sim_F, 10, axis=0)
popF_p90 = np.percentile(pop_sim_F, 90, axis=0)

popM_p10 = np.percentile(pop_sim_M, 10, axis=0)
popM_p90 = np.percentile(pop_sim_M, 90, axis=0)

rows = []

for t, year in enumerate(YEARS):
    for i, age in enumerate(AGE_LABELS):

        rows.append({
            "Sex": "F",
            "Year": year,
            "Age_Group": age,
            "Median": popF_median[t, i],
            "P10": popF_p10[t, i],
            "P90": popF_p90[t, i]
        })
```

```

        rows.append({
            "Sex": "M",
            "Year": year,
            "Age_Group": age,
            "Median": popM_median[t, i],
            "P10": popM_p10[t, i],
            "P90": popM_p90[t, i]
        })

pop_by_age_sex = pd.DataFrame(rows)

pop_by_age_sex.to_excel("forecast_age_sex_with_PI.xlsx", index=False)

print("Файл сохранён: forecast_age_sex_with_PI.xlsx")
display(pop_by_age_sex.head())

# =====
# 3. Компоненты прогноза
# Рождаемость, смертность и миграция выгружаются с медианой и интервальными
# границами.
# =====

def component_PI(model_dict, name, age_labels):
    """
    Формирует таблицу медианы, P10 и P90 для прогнозной компоненты.
    Массив sims имеет размерность: симуляции × годы × возрастные группы.
    """
    sims = model_dict["sims"]

    med = np.median(sims, axis=0)
    p10 = np.percentile(sims, 10, axis=0)
    p90 = np.percentile(sims, 90, axis=0)

    med_df = pd.DataFrame(med, columns=age_labels)
    p10_df = pd.DataFrame(p10, columns=age_labels)
    p90_df = pd.DataFrame(p90, columns=age_labels)

    med_df["Year"] = YEARS
    p10_df["Year"] = YEARS
    p90_df["Year"] = YEARS

    out = med_df.melt(
        id_vars="Year",
        var_name="Age_Group",
        value_name="Median"
    )

    out_p10 = p10_df.melt(
        id_vars="Year",
        var_name="Age_Group",
        value_name="P10"
    )

    out_p90 = p90_df.melt(
        id_vars="Year",
        var_name="Age_Group",
        value_name="P90"
    )

```

```

    )

    out = (
        out
        .merge(out_p10, on=["Year", "Age_Group"], how="left")
        .merge(out_p90, on=["Year", "Age_Group"], how="left")
    )

    out["Component"] = name

    return out

# Для рождаемости используются ASFR_LABELS, поскольку fert_model содержит
# все возрастные группы, представленные в fert_w.
fert_df = component_PI(
    fert_model,
    "Fertility (%)",
    ASFR_LABELS
)

mortF_df = component_PI(
    mort_models["F"],
    "Mortality F (%)",
    AGE_LABELS
)

mortM_df = component_PI(
    mort_models["M"],
    "Mortality M (%)",
    AGE_LABELS
)

migF_df = component_PI(
    mig_models["F"],
    "Migration F (чел.)",
    AGE_LABELS
)

migM_df = component_PI(
    mig_models["M"],
    "Migration M (чел.)",
    AGE_LABELS
)

comp_all = pd.concat(
    [fert_df, mortF_df, mortM_df, migF_df, migM_df],
    ignore_index=True
)

comp_all.to_excel("forecast_components_with_PI.xlsx", index=False)

print("Файл сохранён: forecast_components_with_PI.xlsx")
display(comp_all.head())

# =====
# 4. График общей численности населения
# =====

```

```

plt.figure(figsize=(8, 5))

plt.plot(
    total_df_CAL["Year"],
    total_df_CAL["Median"] / 1e6,
    linewidth=2.5,
    label="Медиана"
)

plt.fill_between(
    total_df_CAL["Year"],
    total_df_CAL["P5"] / 1e6,
    total_df_CAL["P95"] / 1e6,
    alpha=0.25,
    label="90% интервал"
)

plt.fill_between(
    total_df_CAL["Year"],
    total_df_CAL["P12_5"] / 1e6,
    total_df_CAL["P87_5"] / 1e6,
    alpha=0.35,
    label="75% интервал"
)

plt.title("Прогноз общей численности населения Санкт-Петербурга, 2024-2045 гг.")
plt.xlabel("Год")
plt.ylabel("Население, млн человек")
plt.grid(alpha=0.3)
plt.legend()
plt.tight_layout()

plt.savefig("plot_total_forecast.png", dpi=300, bbox_inches="tight")
plt.show()

print("График сохранён: plot_total_forecast.png")

# =====
# 5. Возрастно-половая структура населения в 2045 году
# =====

year_show = 2045

df_show = pop_by_age_sex.query("Year == @year_show").copy()

df_show["Plot_Value"] = np.where(
    df_show["Sex"] == "M",
    -df_show["Median"],
    df_show["Median"]
)

plt.figure(figsize=(7, 8))

for sex in ["M", "F"]:
    sub = df_show[df_show["Sex"] == sex]

    plt.barh(

```

```

        sub["Age_Group"],
        sub["Plot_Value"],
        label=sex
    )

plt.axvline(0, linewidth=1)

plt.title(f"Возрастно-половая структура населения, {year_show} г.")
plt.xlabel("Численность, человек")
plt.ylabel("Возрастная группа")
plt.legend(title="Пол")
plt.tight_layout()

plt.savefig(f"plot_pyramid_{year_show}.png", dpi=300, bbox_inches="tight")
plt.show()

print(f"График сохранён: plot_pyramid_{year_show}.png")

# =====
# 6. График возрастной смертности женщин 65-69 лет
# =====

sub = mortF_df.query("Age_Group == '65-69'").copy()

plt.figure(figsize=(8, 5))

plt.plot(
    sub["Year"],
    sub["Median"],
    linewidth=2.2,
    label="Медиана"
)

plt.fill_between(
    sub["Year"],
    sub["P10"],
    sub["P90"],
    alpha=0.25,
    label="P10-P90"
)

plt.title("Возрастная смертность женщин 65-69 лет, прогноз 2024-2045 гг.")
plt.xlabel("Год")
plt.ylabel("%")
plt.grid(alpha=0.3)
plt.legend()
plt.tight_layout()

plt.savefig("plot_mortality_65_69_F.png", dpi=300, bbox_inches="tight")
plt.show()

print("График сохранён: plot_mortality_65_69_F.png")

# =====
# 7. Тепловая карта возрастных коэффициентов рождаемости ASFR
# =====

```

```

pivot_fert = fert_df.pivot_table(
    index="Age_Group",
    columns="Year",
    values="Median"
)

# Порядок возрастных групп сохраняется в соответствии с данными ASFR.
pivot_fert = pivot_fert.reindex(ASFR_LABELS)

plt.figure(figsize=(10, 5))

plt.imshow(
    pivot_fert.values,
    aspect="auto"
)

plt.xticks(
    ticks=np.arange(len(pivot_fert.columns)),
    labels=pivot_fert.columns,
    rotation=45
)

plt.yticks(
    ticks=np.arange(len(pivot_fert.index)),
    labels=pivot_fert.index
)

plt.colorbar(label="ASFR, %")

plt.title("Возрастные коэффициенты рождаемости, медианная траектория")
plt.xlabel("Год")
plt.ylabel("Возрастная группа")
plt.tight_layout()

plt.savefig("plot_heatmap_fertility.png", dpi=300, bbox_inches="tight")
plt.show()

print("График сохранён: plot_heatmap_fertility.png")

```

```

# === ВЫГРУЗКА КОМПОНЕНТ С ДОВЕРИТЕЛЬНЫМИ ИНТЕРВАЛАМИ (РОБОСТНАЯ) ===
import numpy as np
import pandas as pd

# 0) аккуратно соберём «эталонные» списки меток
AGE_LABELS_ALL = [str(a) for a in AGE_LABELS] # вся
сетка (например, 18)
ASFR_LABELS_ALL = [str(a) for a in getattr(fert_w, "columns", [])] # как
в данных фертильности
REPRO_LABELS_ALL = [str(a) for a in REPRO_AGE_GROUPS] #
твой список (может быть 7 или 8)

def _infer_age_labels(n_cols, prefer="auto"):
    """
    Подбирает подходящие возрастные метки по числу столбцов в sims.
    prefer: 'fert' / 'mortmig' / 'auto'
    """
    # кандидаты с приоритетами
    cand_fert = [ASFR_LABELS_ALL, REPRO_LABELS_ALL, AGE_LABELS_ALL]
    cand_mortmig = [AGE_LABELS_ALL, ASFR_LABELS_ALL, REPRO_LABELS_ALL]

```

```

cands_auto      = cands_fert + cands_mortmig

pool = {"fert": cands_fert, "mortmig": cands_mortmig, "auto":
cands_auto}[prefer]
for lst in pool:
    if lst and len(lst) == n_cols:
        return lst
# fallback - сгенерируем технические ярлыки
return [f"A{i}" for i in range(n_cols)]

def _component_PI(model_dict, years, comp_name, prefer="auto"):
    """
    Делает длинную таблицу: Year, Age_Group, Median, P10, P90, Component
    model_dict['sims'] имеет форму (N_sims, T, A) - как мы делали в частях 2/3.
    """
    sims = model_dict["sims"] # (N, T, A)
    T, A = sims.shape[1], sims.shape[2]
    labels = _infer_age_labels(A, prefer=prefer)

    # перцентили и медиана по симуляциям
    median = np.median(sims, axis=0) # (T, A)
    p10     = np.percentile(sims, 10, axis=0)
    p90     = np.percentile(sims, 90, axis=0)

    # аккуратно «расплавим» без рассинхрона
    out = pd.DataFrame({
        "Year":      np.repeat(years, A),
        "Age_Group": np.tile(labels, T),
        "Median":    median.reshape(-1),
        "P10":       p10.reshape(-1),
        "P90":       p90.reshape(-1),
    })
    out["Component"] = comp_name
    return out

# 1) собираем все компоненты с корректными метками
fert_df = _component_PI(fert_model, YEARS, "Fertility (%)",
prefer="fert")
mortF_df = _component_PI(mort_models["F"], YEARS, "Mortality F (%)",
prefer="mortmig")
mortM_df = _component_PI(mort_models["M"], YEARS, "Mortality M (%)",
prefer="mortmig")
migF_df  = _component_PI(mig_models["F"], YEARS, "Migration F (чел.)",
prefer="mortmig")
migM_df  = _component_PI(mig_models["M"], YEARS, "Migration M (чел.)",
prefer="mortmig")

comp_all = pd.concat([fert_df, mortF_df, mortM_df, migF_df, migM_df],
ignore_index=True)

# 2) сохраняем в Excel
comp_all.to_excel("forecast_components_with_PI.xlsx", index=False)
print("Сохранён: forecast_components_with_PI.xlsx")

```

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

# --- пути к исходным файлам

```

```

hist_path = "/Users/daria/Population_Forecasting_St_Petersburg.xlsx"
fc_path   = "/Users/daria/forecast_total_with_PI.xlsx"

# 1. Историческая численность населения за 1990–2023 гг.
# Рассчитывается как сумма по всем возрастно-половым группам.
hist = pd.read_excel(hist_path, sheet_name="Base Population Data")
hist = (hist.groupby("Year", as_index=False)["Population"].sum()
        .query("Year>=1990 & Year<=2023")
        .rename(columns={"Population": "Value"}))
hist["Year"] = hist["Year"].astype(int)

# 2. Прогнозная численность населения за 2024–2045 гг.
fc = pd.read_excel(fc_path)

for c in ["Year", "Median", "P5", "P95", "P12_5", "P87_5"]:
    if c in fc.columns:
        fc[c] = pd.to_numeric(fc[c], errors="coerce")

fc = fc.sort_values("Year").query("Year>=2024").copy()

# 3. Стыковка исторического и прогнозного рядов.
# В прогнозный ряд добавляется базовый 2023 год,
# для которого интервальные границы совпадают с фактической численностью.
base_2023 = float(hist.loc[hist["Year"]==2023, "Value"])

fc0 = pd.DataFrame([
    "Year": 2023,
    "Median": base_2023,
    "P5": base_2023,
    "P95": base_2023,
    "P12_5": base_2023,
    "P87_5": base_2023
])

fc = pd.concat([fc0, fc], ignore_index=True)

# 4. Проверка корректности интервальных границ.
def _swap_if_wrong(df, a, b):
    bad = df[a] > df[b]
    if bad.any():
        df.loc[bad, [a, b]] = df.loc[bad, [b, a]].to_numpy()
    return df

fc = _swap_if_wrong(fc, "P5", "P95")
fc = _swap_if_wrong(fc, "P12_5", "P87_5")

# Медианная траектория ограничивается границами 90% интервала.
fc["Median"] = np.clip(
    fc["Median"].to_numpy(),
    fc["P5"].to_numpy(),
    fc["P95"].to_numpy()
)

# 5. Построение графика динамики и прогноза численности населения.
plt.figure(figsize=(11, 5.2))

# Историческая динамика.
plt.plot(
    hist["Year"],

```

```

    hist["Value"] / 1e6,
    color="#6F6F6F",
    lw=2.2,
    label="Фактическая численность (1990–2023)",
    zorder=3
)

# 90% предиктивный интервал.
plt.fill_between(
    fc["Year"],
    fc["P5"] / 1e6,
    fc["P95"] / 1e6,
    color="#E6E6E6",
    alpha=1.0,
    label="90% ДИ",
    zorder=1
)

# 75% предиктивный интервал.
plt.fill_between(
    fc["Year"],
    fc["P12_5"] / 1e6,
    fc["P87_5"] / 1e6,
    color="#C9C9C9",
    alpha=1.0,
    label="75% ДИ",
    zorder=2
)

# Медианная прогнозная траектория.
plt.plot(
    fc["Year"],
    fc["Median"] / 1e6,
    color="#000000",
    lw=2.6,
    label="Прогноз (медиана)",
    zorder=4
)

# Обозначение базового года прогноза.
plt.axvline(2023, color="#000000", lw=1, ls="--", alpha=0.5)
plt.text(
    2023 + 0.2,
    plt.ylim()[0] + 0.02,
    "Базовый год (2023)",
    fontsize=9,
    alpha=0.7
)

plt.title("Динамика и прогноз численности населения Санкт-Петербурга, 1990–2045")
plt.xlabel("Год")
plt.ylabel("Млн человек")
plt.grid(alpha=0.3)
plt.legend()
plt.tight_layout()

plt.savefig("Population_Forecast_1990_2045.png", dpi=300, bbox_inches="tight")
plt.show()

```

```

import pandas as pd
import matplotlib.pyplot as plt

# === 1. Загрузка данных прогноза ===
df = pd.read_excel("forecast_total_with_PI.xlsx")

for c in ["Year", "Median", "P5", "P95", "P12_5", "P87_5"]:
    df[c] = pd.to_numeric(df[c], errors="coerce")

# === 2. Прогноз Росстата ===
rosstat_data = {
    "Year": [2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034,
             2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045,
             2046],
    "Rosstat": [5607300, 5608561, 5608305, 5606543, 5603601, 5600318, 5596676,
                5592539, 5588100, 5583500, 5579400, 5575600, 5571836, 5568200,
                5564900, 5561400, 5557500, 5552975, 5547900, 5542200, 5535900,
                5528900, 5521143]
}

rosstat = pd.DataFrame(rosstat_data)

# === 3. Годы для подписей значений ===
years_our = [2025, 2030, 2036, 2041, 2045]
years_ros = [2025, 2030, 2036, 2041, 2046]

# === 4. Построение графика ===
plt.figure(figsize=(10, 5.5))

# 90% предиктивный интервал.
plt.fill_between(
    df["Year"],
    df["P5"] / 1e6,
    df["P95"] / 1e6,
    alpha=1.0,
    color="#E6E6E6",
    label="90% интервал"
)

# 75% предиктивный интервал.
plt.fill_between(
    df["Year"],
    df["P12_5"] / 1e6,
    df["P87_5"] / 1e6,
    alpha=1.0,
    color="#C9C9C9",
    label="75% интервал"
)

# Медианная траектория авторского прогноза.
plt.plot(
    df["Year"],
    df["Median"] / 1e6,
    lw=2.5,
    color="#000000",
    label="Прогноз, медиана"
)

```

```

# Средний вариант прогноза Росстата.
plt.plot(
    rosstat["Year"],
    rosstat["Rosstat"] / 1e6,
    color="#6F6F6F",
    lw=2,
    ls="--",
    marker="o",
    label="Прогноз Росстата"
)

# === 5. Подписи значений ===

# Значения медианной траектории авторского прогноза.
for y in years_our:
    val = df.loc[df["Year"] == y, "Median"].values
    if len(val) > 0:
        plt.text(
            Y,
            val[0] / 1e6 + 0.02,
            f"{val[0] / 1e6:.2f}",
            ha="center",
            va="bottom",
            fontsize=9,
            color="#000000",
            bbox=dict(boxstyle="round,pad=0.25", fc="white", alpha=0.7, lw=0)
        )

# Значения прогноза Росстата.
for y in years_ros:
    val = rosstat.loc[rosstat["Year"] == y, "Rosstat"].values
    if len(val) > 0:
        plt.text(
            Y,
            val[0] / 1e6 - 0.04,
            f"{val[0] / 1e6:.2f}",
            ha="center",
            va="top",
            fontsize=9,
            color="#4F4F4F",
            bbox=dict(boxstyle="round,pad=0.25", fc="white", alpha=0.7, lw=0)
        )

# === 6. Оформление графика ===
plt.xlabel("Год")
plt.ylabel("Млн человек")
plt.grid(alpha=0.3)
plt.legend()
plt.tight_layout()

# === 7. Сохранение графика ===
plt.savefig("forecast_total_vs_rosstat_2024_2045_labeled.png", dpi=300,
            bbox_inches="tight")
plt.show()

```

```

import pandas as pd
import matplotlib.pyplot as plt
from matplotlib.ticker import FuncFormatter

```

```

# === 1. Загрузка и подготовка данных ===
PATH = "forecast_age_sex_with_PI.xlsx"
df = pd.read_excel(PATH)

# Нормализация имён колонок.
cols = {c.lower(): c for c in df.columns}

def pick(*names):
    for n in names:
        if n in cols:
            return cols[n]
    raise KeyError(f"Отсутствуют необходимые колонки: {names}")

col_year = pick("year")
col_age = pick("age_group", "age", "agegroup")
col_sex = pick("sex", "gender")
col_val = pick("median", "value", "population")

df = df[[col_year, col_age, col_sex, col_val]].rename(
    columns={
        col_year: "Year",
        col_age: "Age_Group",
        col_sex: "Sex",
        col_val: "Value"
    }
).copy()

# Приведение типов данных.
df["Year"] = pd.to_numeric(df["Year"], errors="coerce")
df["Value"] = pd.to_numeric(df["Value"], errors="coerce")
df["Age_Group"] = df["Age_Group"].astype(str)
df["Sex"] = df["Sex"].astype(str)

# Функция сортировки возрастных групп.
def age_key(g):
    g = str(g).strip()
    if g.endswith("+"):
        return (int(g[:-1]), 999)
    a, b = g.split("-")
    return (int(a), int(b))

# Порядок возрастных групп.
ages = sorted(df["Age_Group"].dropna().unique(), key=age_key)

# Годы для построения возрастно-половых пирамид.
YEARS_SHOW = [2030, 2035, 2040, 2045]

# === 2. Формирование рядов по полу для каждого года ===
def get_FM_by_year(y):
    sub = df[df["Year"] == y].copy()

    F = (
        sub[sub["Sex"].str.upper().str.startswith("F")]
        .set_index("Age_Group")["Value"]
        .reindex(ages)
        .fillna(0.0)
    )

    M = (

```

```

        sub[sub["Sex"].str.upper().str.startswith("M")]
        .set_index("Age_Group")["Value"]
        .reindex(ages)
        .fillna(0.0)
    )

    return F, M

# Единая шкала для всех пирамид.
max_bin = 0.0
cache = {}

for y in YEARS_SHOW:
    F, M = get_FM_by_year(y)
    cache[y] = (F, M)
    max_bin = max(max_bin, F.max(), M.max())

max_bin = float(max_bin) * 1.10

# Форматирование оси X в тыс. человек.
fmt_thousands = FuncFormatter(lambda x, pos: f"{abs(x)/1e3:.0f}")

# === 3. Построение возрастно-половых пирамид ===
fig, axes = plt.subplots(
    nrows=len(YEARS_SHOW),
    ncols=1,
    figsize=(10, 16),
    sharex=True
)

COLOR_M = "#BFBFBF"
COLOR_F = "#6F6F6F"

def plot_pyramid(ax, year, F, M):
    # Мужчины отображаются слева, женщины - справа.
    ax.barh(
        ages,
        -M.values,
        color=COLOR_M,
        edgecolor="white",
        height=0.9,
        label="Мужчины"
    )

    ax.barh(
        ages,
        F.values,
        color=COLOR_F,
        edgecolor="white",
        height=0.9,
        label="Женщины"
    )

    # Центральная ось.
    ax.axvline(0, color="#000000", lw=1, alpha=0.5)

    # Заголовок панели.
    ax.set_title(f"{year} год", loc="left", fontsize=12, pad=6)

```

```

# Единая шкала по оси X.
ax.set_xlim(-max_bin, max_bin)

# Подписи сторон пирамиды.
ax.text(
    -max_bin * 0.96,
    ages[-1],
    "М",
    fontsize=12,
    color="#000000",
    ha="left",
    va="bottom",
    fontweight="bold",
    bbox=dict(boxstyle="round,pad=0.2", fc="white", lw=0, alpha=0.7)
)

ax.text(
    max_bin * 0.96,
    ages[-1],
    "Ж",
    fontsize=12,
    color="#000000",
    ha="right",
    va="bottom",
    fontweight="bold",
    bbox=dict(boxstyle="round,pad=0.2", fc="white", lw=0, alpha=0.7)
)

ax.xaxis.set_major_formatter(fmt_thousands)
ax.grid(axis="x", alpha=0.25)
ax.set_ylabel("Возрастная группа")

for ax, y in zip(axes, YEARS_SHOW):
    F, M = cache[y]
    plot_pyramid(ax, y, F, M)

# Подпись общей оси X.
axes[-1].set_xlabel("Численность по полу и возрасту, тыс. человек", fontsize=11)

# Легенда.
handles = [
    plt.Line2D([0], [0], color=COLOR_M, lw=8),
    plt.Line2D([0], [0], color=COLOR_F, lw=8),
]

axes[0].legend(
    handles,
    ["Мужчины", "Женщины"],
    loc="upper left",
    frameon=True
)

plt.tight_layout()
plt.savefig("pyramids_2030_2035_2040_2045.png", dpi=300, bbox_inches="tight")
plt.show()

```

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

```

```

from matplotlib.ticker import FuncFormatter

# === 1. Загрузка и подготовка данных ===
PATH = "forecast_age_sex_with_PI.xlsx"
df = pd.read_excel(PATH)

# Нормализация названий колонок.
cols = {c.lower(): c for c in df.columns}

def pick(*names):
    for n in names:
        if n in cols:
            return cols[n]
        raise KeyError(f"Отсутствуют необходимые колонки: {names}")

col_year = pick("year")
col_age = pick("age_group", "age", "agegroup")
col_sex = pick("sex", "gender")
col_val = pick("median", "value", "population")

df = df[[col_year, col_age, col_sex, col_val]].rename(
    columns={
        col_year: "Year",
        col_age: "Age_Group",
        col_sex: "Sex",
        col_val: "Value"
    }
).copy()

df["Year"] = pd.to_numeric(df["Year"], errors="coerce")
df["Value"] = pd.to_numeric(df["Value"], errors="coerce")
df["Age_Group"] = df["Age_Group"].astype(str)
df["Sex"] = df["Sex"].astype(str)

# Функция сортировки возрастных групп.
def age_key(g):
    g = str(g).strip()
    if g.endswith("+"):
        return (int(g[:-1]), 999)
    a, b = g.split("-")
    return (int(a), int(b))

ages = sorted(df["Age_Group"].unique(), key=age_key)

# Годы для построения возрастно-половых пирамид.
YEARS_SHOW = [2030, 2035, 2040, 2045]

# === 2. Формирование рядов по полу для каждого года ===
def get_FM_by_year(y):
    sub = df[df["Year"] == y].copy()

    F = (
        sub[sub["Sex"].str.upper().str.startswith("F")]
        .set_index("Age_Group")["Value"]
        .reindex(ages)
        .fillna(0.0)
    )

    M = (

```

```

        sub[sub["Sex"].str.upper().str.startswith("M")]
        .set_index("Age_Group")["Value"]
        .reindex(ages)
        .fillna(0.0)
    )

    return F, M

# Единая шкала для всех пирамид.
max_bin = 0.0
cache = {}

for y in YEARS_SHOW:
    F, M = get_FM_by_year(y)
    cache[y] = (F, M)
    max_bin = max(max_bin, F.max(), M.max())

max_bin = float(max_bin) * 1.10

fmt_thousands = FuncFormatter(lambda x, pos: f"{abs(x)/1e3:.0f}")

# === 3. Построение возрастно-половых пирамид ===
fig, axes = plt.subplots(
    nrows=1,
    ncols=len(YEARS_SHOW),
    figsize=(16, 6),
    sharey=True
)

COLOR_M = "#BFBFBF"
COLOR_F = "#6F6F6F"

def plot_pyramid(ax, year, F, M):
    # Мужчины отображаются слева, женщины — справа.
    ax.barh(ages, -M.values, color=COLOR_M, edgecolor="white", height=0.9)
    ax.barh(ages, F.values, color=COLOR_F, edgecolor="white", height=0.9)

    ax.axvline(0, color="#000000", lw=1, alpha=0.5)
    ax.set_xlim(-max_bin, max_bin)
    ax.set_title(f"{year}", fontsize=12, pad=8)

    # Обозначения сторон пирамиды.
    ax.text(
        -max_bin * 0.95,
        ages[-1],
        "М",
        fontsize=12,
        color="#000000",
        ha="left",
        va="bottom",
        alpha=0.9,
        bbox=dict(boxstyle="round,pad=0.2", fc="white", lw=0, alpha=0.7)
    )

    ax.text(
        max_bin * 0.95,
        ages[-1],
        "Ж",
        fontsize=12,

```

```

        color="#000000",
        ha="right",
        va="bottom",
        alpha=0.9,
        bbox=dict(boxstyle="round,pad=0.2", fc="white", lw=0, alpha=0.7)
    )

    ax.xaxis.set_major_formatter(fmt_thousands)
    ax.grid(axis="x", alpha=0.25)

for ax, y in zip(axes, YEARS_SHOW):
    F, M = cache[y]
    plot_pyramid(ax, y, F, M)

# === 4. Общее оформление ===
axes[0].set_ylabel("Возрастная группа", fontsize=11)
fig.text(0.5, 0.00, "Численность, тыс. человек", ha="center", fontsize=11)
fig.suptitle("Возрастно-половая структура населения Санкт-Петербурга, 2030-2045", fontsize=14, y=1.02)

# === 5. Легенда ===
handles = [
    plt.Line2D([0], [0], color=COLOR_M, lw=8),
    plt.Line2D([0], [0], color=COLOR_F, lw=8),
]
axes[-1].legend(handles, ["Мужчины", "Женщины"], loc="lower right",
frameon=True)

plt.tight_layout()
plt.savefig("pyramids_horizontal_2030_2035_2040_2045.png", dpi=300,
bbox_inches="tight")
plt.show()

```

```

# =====
# Выгрузка чистой миграции, применённой в прогнозной модели
# с учётом эмпирических возрастных весов, ограничений по группам
# и уровневой добавки к чистой миграции
# =====

import numpy as np
import pandas as pd

# -----
# 1. Векторизованное распределение миграции по возрасту
# Аналог функции shape_migration_sex для массива  $N \times T \times A$ .
# -----

def shape_migration_batch(mig_raw, W, cap=MIG_GROUP_CAP):
    """
    mig_raw: массив  $N \times T \times A$  с траекториями чистой миграции из PCA+ARIMA.
    W: массив  $A$  с эмпирическими возрастными весами по полу.
    """
    v = mig_raw.astype(float)

    # Применение эмпирических возрастных весов.
    v_w = v * W[None, None, :]

    # Общий объём миграции до и после взвешивания.
    s = v.sum(axis=2, keepdims=True)

```

```

sw = v_w.sum(axis=2, keepdims=True)

# Перенормировка для сохранения общего объёма миграции.
ratio = np.divide(
    s,
    sw,
    out=np.zeros_like(s),
    where=np.abs(sw) >= 1e-12
)

adj = v_w * ratio

# Применение ограничения по возрастной группе.
return np.clip(adj, -cap, cap)

# -----
# 2. Исходные стохастические траектории миграции из моделей
# -----

migF_raw = mig_models["F"]["sims"] # N_SIMS × len(YEARS) × len(AGE_LABELS)
migM_raw = mig_models["M"]["sims"]

# -----
# 3. Распределение миграции по возрасту и добавление уровневой корректировки
# -----

migF_shaped = shape_migration_batch(
    migF_raw,
    W=wF,
    cap=MIG_GROUP_CAP
)

migM_shaped = shape_migration_batch(
    migM_raw,
    W=wM,
    cap=MIG_GROUP_CAP
)

bumpF = (
    MIG_NET_BUMP *
    MIG_BUMP_SEX_SPLIT[0]
) * bump_weights.values

bumpM = (
    MIG_NET_BUMP *
    MIG_BUMP_SEX_SPLIT[1]
) * bump_weights.values

migF_used = migF_shaped + bumpF[None, None, :]
migM_used = migM_shaped + bumpM[None, None, :]

# -----
# 4. Итоговые показатели чистой миграции по годам
# -----

migF_total = migF_used.sum(axis=2)

```

```

migM_total = migM_used.sum(axis=2)
migALL_total = migF_total + migM_total

def summarize_PI(arr_NT):
    """
    arr_NT: массив N × T.
    Возвращает медиану и интервальные границы по годам.
    """
    return {
        "Median": np.median(arr_NT, axis=0),
        "P5": np.percentile(arr_NT, 5, axis=0),
        "P95": np.percentile(arr_NT, 95, axis=0),
        "P12_5": np.percentile(arr_NT, 12.5, axis=0),
        "P87_5": np.percentile(arr_NT, 87.5, axis=0),
    }

S_all = summarize_PI(migALL_total)
S_f = summarize_PI(migF_total)
S_m = summarize_PI(migM_total)

# Число прогнозных лет определяется по массиву результатов.
T_forecast = migALL_total.shape[1]

# Прогнозные годы приводятся к длине массива результатов.
YEARS_USED = np.array(YEARS[:T_forecast])

print("Количество лет в YEARS:", len(YEARS))
print("Количество лет в прогнозном массиве:", T_forecast)
print("Используемый прогнозный период:", YEARS_USED[0], "-", YEARS_USED[-1])

mig_year_df = pd.DataFrame({
    "Year": YEARS_USED,

    "NetMig_Median": S_all["Median"],
    "NetMig_P5": S_all["P5"],
    "NetMig_P95": S_all["P95"],
    "NetMig_P12_5": S_all["P12_5"],
    "NetMig_P87_5": S_all["P87_5"],

    "NetMigF_Median": S_f["Median"],
    "NetMigF_P5": S_f["P5"],
    "NetMigF_P95": S_f["P95"],

    "NetMigM_Median": S_m["Median"],
    "NetMigM_P5": S_m["P5"],
    "NetMigM_P95": S_m["P95"],
})

mig_year_df.to_excel(
    "forecast_net_migration_used_PI.xlsx",
    index=False
)

print("Файл сохранён: forecast_net_migration_used_PI.xlsx")
display(mig_year_df.head())
display(mig_year_df.tail())

```

```

# -----
# 5. Возрастно-половой состав применённой чистой миграции
# -----

def age_profile_long(mig_used_NTA, sex_label):
    """
    mig_used_NTA: массив  $N \times T \times A$ .
    Возвращает длинную таблицу  $Year \times Age\_Group \times Sex$ .
    """
    med = np.median(mig_used_NTA, axis=0)
    p5 = np.percentile(mig_used_NTA, 5, axis=0)
    p95 = np.percentile(mig_used_NTA, 95, axis=0)

    T = med.shape[0]
    A_n = med.shape[1]

    years_used = YEARS_USED[:T]
    age_labels_used = list(AGE_LABELS)[:A_n]

    out = pd.DataFrame({
        "Year": np.repeat(years_used, A_n),
        "Age_Group": np.tile(age_labels_used, T),
        "Sex": sex_label,
        "Median": med.reshape(-1),
        "P5": p5.reshape(-1),
        "P95": p95.reshape(-1),
    })

    return out

mig_age_df = pd.concat(
    [
        age_profile_long(migF_used, "F"),
        age_profile_long(migM_used, "M")
    ],
    ignore_index=True
)

mig_age_df.to_excel(
    "forecast_net_migration_used_by_age.xlsx",
    index=False
)

print("Файл сохранён: forecast_net_migration_used_by_age.xlsx")
display(mig_age_df.head())

```

```

import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_excel("forecast_net_migration_used_PI.xlsx")

# --- проверка корректности интервальных границ ---
# Квантили должны быть упорядочены монотонно.
assert (df["NetMig_P5"] <= df["NetMig_P12_5"]).all()
assert (df["NetMig_P12_5"] <= df["NetMig_Median"]).all()
assert (df["NetMig_Median"] <= df["NetMig_P87_5"]).all()
assert (df["NetMig_P87_5"] <= df["NetMig_P95"]).all()

```

```

print("Квантили упорядочены корректно")

# --- контрольный вывод первых строк в тыс. человек ---
tmp = df.loc[
    :,
    ["Year", "NetMig_Median", "NetMig_P5", "NetMig_P95", "NetMig_P12_5",
    "NetMig_P87_5"]
].copy()

for c in tmp.columns[1:]:
    tmp[c] = tmp[c] / 1000

print(tmp.head(5))

# --- подготовка данных для графика ---
SCALE = 1000.0

x = df["Year"].values
med = df["NetMig_Median"].values / SCALE
p5 = df["NetMig_P5"].values / SCALE
p95 = df["NetMig_P95"].values / SCALE
p12 = df["NetMig_P12_5"].values / SCALE
p87 = df["NetMig_P87_5"].values / SCALE

# --- построение графика чистой миграции ---
plt.figure(figsize=(10, 5))

# 90% предиктивный интервал.
plt.fill_between(
    x,
    p5,
    p95,
    alpha=1.0,
    color="#E6E6E6",
    label="90% интервал",
    zorder=1
)

# 75% предиктивный интервал.
plt.fill_between(
    x,
    p12,
    p87,
    alpha=1.0,
    color="#C9C9C9",
    label="75% интервал",
    zorder=2
)

# Медианная траектория.
plt.plot(
    x,
    med,
    linewidth=2,
    color="#000000",
    label="Медиана",
    zorder=3
)

```

```

# Нулевая линия.
plt.axhline(
    0,
    color="#6F6F6F",
    linestyle="--",
    linewidth=1
)

plt.xlabel("Год")
plt.ylabel("Чистая миграция, тыс. человек")
plt.title("Прогноз чистой миграции: медиана и предиктивные интервалы")
plt.legend()
plt.tight_layout()

plt.savefig("net_migration_used_PI_clean.png", dpi=300, bbox_inches="tight")
plt.show()

print("График сохранён: net_migration_used_PI_clean.png")

```

```

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

# =====
# 1. Подготовка индексов
# =====
# Требуются результаты предыдущих частей:
# YEARS, AGE_LABELS, ASFR_LABELS, REPRO_AGE_GROUPS,
# pop_sim_F, pop_sim_M,
# mort_models, mig_models, fert_model,
# base_F, base_M,
# shape_migration_sex, surv_prob_exp_by_age,
# MIG_NET_BUMP, MIG_BUMP_SEX_SPLIT, bump_weights, MIG_GROUP_CAP.

YEARS = list(YEARS)
N = pop_sim_F.shape[0]
T = len(YEARS)

AGE_LABELS = [str(a) for a in AGE_LABELS]
ASFR_LABELS = [str(a) for a in ASFR_LABELS]
REPRO_AGE_GROUPS = [str(a) for a in REPRO_AGE_GROUPS]

# Сопоставление репродуктивных возрастных групп в структуре населения и в ASFR.
repro_ok = [g for g in REPRO_AGE_GROUPS if (g in AGE_LABELS and g in
ASFR_LABELS)]

if len(repro_ok) < 5:
    raise ValueError(f"Недостаточно совпадающих репродуктивных групп:
{repro_ok}")

idx_pop_repro = np.array([AGE_LABELS.index(g) for g in repro_ok], dtype=int)
idx_asfr_repro = np.array([ASFR_LABELS.index(g) for g in repro_ok], dtype=int)

# Возрастные векторы уровневой добавки к чистой миграции.
bumpF_vec = (MIG_NET_BUMP * MIG_BUMP_SEX_SPLIT[0]) * bump_weights.values
bumpM_vec = (MIG_NET_BUMP * MIG_BUMP_SEX_SPLIT[1]) * bump_weights.values

```

```

# =====
# 2. Расчёт компонент по стохастическим траекториям
# =====

births = np.zeros((N, T), dtype=float)
deaths = np.zeros((N, T), dtype=float)
netmig = np.zeros((N, T), dtype=float)

for d in range(N):
    for t in range(T):
        # Стартовая численность населения для расчёта смертей в году t.
        if t == 0:
            popF_start = base_F
            popM_start = base_M
        else:
            popF_start = pop_sim_F[d, t-1, :]
            popM_start = pop_sim_M[d, t-1, :]

        # Расчёт числа умерших по возрастно-половым коэффициентам смертности.
        mortF_t = mort_models["F"]["sims"][d, t, :]
        mortM_t = mort_models["M"]["sims"][d, t, :]

        survF = surv_prob_exp_by_age(mortF_t)
        survM = surv_prob_exp_by_age(mortM_t)

        deaths[d, t] = (
            (popF_start * (1 - survF)).sum() +
            (popM_start * (1 - survM)).sum()
        )

        # Расчёт чистой миграции, применённой в прогнозной модели.
        migF_t = mig_models["F"]["sims"][d, t, :]
        migM_t = mig_models["M"]["sims"][d, t, :]

        migF_used_vec = shape_migration_sex(migF_t, "F", cap=MIG_GROUP_CAP) +
        bumpF_vec
        migM_used_vec = shape_migration_sex(migM_t, "M", cap=MIG_GROUP_CAP) +
        bumpM_vec

        netmig[d, t] = migF_used_vec.sum() + migM_used_vec.sum()

        # Расчёт числа рождений по возрастным коэффициентам рождаемости.
        # Добавление новорождённых влияет только на группу 0-4,
        # поэтому репродуктивные возрастные группы не искажаются.
        asfr_t = fert_model["sims"][d, t, :]
        popF_repro = pop_sim_F[d, t, idx_pop_repro]
        asfr_repro = asfr_t[idx_asfr_repro]

        births[d, t] = (popF_repro * asfr_repro).sum() / 1000.0

print("Компоненты рассчитаны: births, deaths, netmig")

# =====
# 3. Формирование таблиц процентилей
# =====

def summarize(arr_NT):
    return pd.DataFrame({

```

```

        "Year": YEARS,
        "Median": np.median(arr_NT, axis=0),
        "P5": np.percentile(arr_NT, 5, axis=0),
        "P95": np.percentile(arr_NT, 95, axis=0),
        "P12_5": np.percentile(arr_NT, 12.5, axis=0),
        "P87_5": np.percentile(arr_NT, 87.5, axis=0),
    })

births_df = summarize(births)
deaths_df = summarize(deaths)
netmig_df = summarize(netmig)

births_df.to_excel("forecast_births_PI.xlsx", index=False)
deaths_df.to_excel("forecast_deaths_PI.xlsx", index=False)
netmig_df.to_excel("forecast_netmig_used_PI.xlsx", index=False)

print("Сохранены таблицы:")
print("forecast_births_PI.xlsx")
print("forecast_deaths_PI.xlsx")
print("forecast_netmig_used_PI.xlsx")

# =====
# 4. Построение графиков компонент, тыс. человек
# =====

SCALE = 1000.0

def plot_component(df, title, filename):
    x = df["Year"].values
    med = df["Median"].values / SCALE
    p5 = df["P5"].values / SCALE
    p95 = df["P95"].values / SCALE
    p12 = df["P12_5"].values / SCALE
    p87 = df["P87_5"].values / SCALE

    plt.figure(figsize=(10, 5))

    # 90% предиктивный интервал.
    plt.fill_between(
        x,
        p5,
        p95,
        alpha=1.0,
        color="#E6E6E6",
        label="90% интервал"
    )

    # 75% предиктивный интервал.
    plt.fill_between(
        x,
        p12,
        p87,
        alpha=1.0,
        color="#C9C9C9",
        label="75% интервал"
    )

```

```

# Медианная траектория.
plt.plot(
    x,
    med,
    linewidth=2,
    color="#000000",
    label="Медиана"
)

# Нулевая линия.
plt.axhline(
    0,
    color="#6F6F6F",
    linestyle="--",
    linewidth=1
)

plt.xlabel("Год")
plt.ylabel("Тыс. человек")
plt.title(title)
plt.legend()
plt.tight_layout()

plt.savefig(filename, dpi=300, bbox_inches="tight")
plt.show()

print(f"График сохранён: {filename}")

plot_component(
    births_df,
    "Прогноз числа рождений: медиана и предиктивные интервалы",
    "plot_births_PI.png"
)

plot_component(
    deaths_df,
    "Прогноз числа умерших: медиана и предиктивные интервалы",
    "plot_deaths_PI.png"
)

plot_component(
    netmig_df,
    "Прогноз чистой миграции: медиана и предиктивные интервалы",
    "plot_netmig_used_PI.png"
)

# Расчёт и визуализация естественного прироста.
natinc = births - deaths
natinc_df = summarize(natinc)

plot_component(
    natinc_df,
    "Прогноз естественного прироста: медиана и предиктивные интервалы",
    "plot_natural_increase_PI.png"
)

```

```

import numpy as np
import pandas as pd

```

```

import matplotlib.pyplot as plt

# --- проверка наличия необходимых объектов ---
required = [
    "YEARS", "N_SIMS", "AGE_LABELS", "A", "pop_wide",
    "mort_models", "mig_models", "fert_model",
    "surv_prob_exp_by_age", "shape_migration_sex", "births_from_asfr",
    "MIG_GROUP_CAP", "MIG_NET_BUMP", "MIG_BUMP_SEX_SPLIT", "bump_weights", "SRB"
]

missing = [x for x in required if x not in globals()]

if missing:
    raise NameError(
        f"Не найдены переменные: {missing}. Необходимо предварительно выполнить ячейки с моделью."
    )

N = int(N_SIMS)
T = len(YEARS)

p_male = SRB / (1 + SRB)
p_fem = 1 - p_male

base_F = pop_wide["F"].astype(float).values.copy()
base_M = pop_wide["M"].astype(float).values.copy()

# --- массивы для хранения компонент ---
births_tot = np.zeros((N, T)) # число рождений
deaths_tot = np.zeros((N, T)) # число умерших
mig_tot     = np.zeros((N, T)) # применённая чистая миграция
nat_inc     = np.zeros((N, T)) # естественный прирост
delta_pop   = np.zeros((N, T)) # годовое изменение численности населения
cap_loss    = np.zeros((N, T)) # изменение миграционного объёма после
# применения ограничений

for d in range(N):
    popF = base_F.copy()
    popM = base_M.copy()
    prev_total = popF.sum() + popM.sum()

    for t in range(T):
        mortF = mort_models["F"]["sims"][d, t, :]
        mortM = mort_models["M"]["sims"][d, t, :]
        migF_raw = mig_models["F"]["sims"][d, t, :]
        migM_raw = mig_models["M"]["sims"][d, t, :]
        asfr = fert_model["sims"][d, t, :]

        # 1. Смертность рассчитывается на входной численности года.
        survF = surv_prob_exp_by_age(mortF)
        survM = surv_prob_exp_by_age(mortM)

        D = (
            (popF * (1 - survF)).sum() +
            (popM * (1 - survM)).sum()
        )

        # 2. Дожитие и годовой переход между возрастными группами.
        popF = A @ (popF * survF)

```

```

popM = A @ (popM * survM)

# 3. Миграция с учётом возрастного распределения и ограничения по
группе.
migF_vec = shape_migration_sex(migF_raw, "F", cap=MIG_GROUP_CAP)
migM_vec = shape_migration_sex(migM_raw, "M", cap=MIG_GROUP_CAP)

popF += migF_vec
popM += migM_vec

# 4. Уровневая добавка к чистой миграции.
bumpF = (MIG_NET_BUMP * MIG_BUMP_SEX_SPLIT[0]) * bump_weights.values
bumpM = (MIG_NET_BUMP * MIG_BUMP_SEX_SPLIT[1]) * bump_weights.values

popF += bumpF
popM += bumpM

Mig_applied = migF_vec.sum() + migM_vec.sum() + bumpF.sum() +
bumpM.sum()
Mig_raw_plus_bump = migF_raw.sum() + migM_raw.sum() + bumpF.sum() +
bumpM.sum()

cap_loss[d, t] = Mig_applied - Mig_raw_plus_bump

# 5. Рождения и распределение новорождённых по полу.
B = births_from_asfr(popF, asfr)

popF[0] += B * p_fem
popM[0] += B * p_male

total = popF.sum() + popM.sum()

births_tot[d, t] = B
deaths_tot[d, t] = D
mig_tot[d, t] = Mig_applied
nat_inc[d, t] = B - D
delta_pop[d, t] = total - prev_total

prev_total = total

# --- проверка балансового соотношения ---
diff = delta_pop - (nat_inc + mig_tot)

print(
    "Проверка баланса:  $\max|\Delta P - ((B-D) + Mig)| =$ ",
    float(np.max(np.abs(diff)))
)

print("Изменение миграционного объёма после применения ограничений, медиана по
годам, тыс. человек:")
print(pd.Series(np.median(cap_loss, axis=0) / 1000,
index=YEARS).round(2).head(), "...")

# --- функция построения графика с предиктивными интервалами ---
def plot_pi(sims, title, ylabel, fname):
    yrs = np.array(YEARS)

    med = np.median(sims, axis=0)

```

```

p5, p95 = np.percentile(sims, [5, 95], axis=0)
p12, p88 = np.percentile(sims, [12.5, 87.5], axis=0)

plt.figure(figsize=(10, 5))

# 90% предиктивный интервал.
plt.fill_between(
    yrs,
    p5,
    p95,
    alpha=1.0,
    color="#E6E6E6",
    label="90% интервал"
)

# 75% предиктивный интервал.
plt.fill_between(
    yrs,
    p12,
    p88,
    alpha=1.0,
    color="#C9C9C9",
    label="75% интервал"
)

# Медианная траектория.
plt.plot(
    yrs,
    med,
    linewidth=2,
    color="#000000",
    label="Медиана"
)

# Нулевая линия.
plt.axhline(
    0,
    color="#6F6F6F",
    linestyle="--",
    linewidth=1
)

plt.title(title)
plt.xlabel("Год")
plt.ylabel(ylabel)
plt.legend()
plt.tight_layout()

plt.savefig(fname, dpi=300, bbox_inches="tight")
plt.show()

print("График сохранён:", fname)

# --- ВИЗУАЛИЗАЦИЯ КОМПОНЕНТ В ТЫС. ЧЕЛОВЕК ---
plot_pi(
    births_tot / 1000,
    "Прогноз числа рождений: медиана и предиктивные интервалы",
    "Тыс. человек",

```

```

    "comp_births_PI.png"
)

plot_pi(
    deaths_tot / 1000,
    "Прогноз числа умерших: медиана и предиктивные интервалы",
    "Тыс. человек",
    "comp_deaths_PI.png"
)

plot_pi(
    mig_tot / 1000,
    "Прогноз чистой миграции: медиана и предиктивные интервалы",
    "Тыс. человек",
    "comp_netmig_PI.png"
)

plot_pi(
    nat_inc / 1000,
    "Прогноз естественного прироста: медиана и предиктивные интервалы",
    "Тыс. человек",
    "comp_natinc_PI.png"
)

plot_pi(
    delta_pop / 1000,
    "Прогноз годового изменения численности населения: медиана и предиктивные интервалы",
    "Тыс. человек",
    "comp_deltaP_PI.png"
)

```

```

def summarize_component(arr_NT):
    return pd.DataFrame({
        "Year": YEARS,
        "Median": np.median(arr_NT, axis=0),
        "P5": np.percentile(arr_NT, 5, axis=0),
        "P95": np.percentile(arr_NT, 95, axis=0),
        "P12_5": np.percentile(arr_NT, 12.5, axis=0),
        "P87_5": np.percentile(arr_NT, 87.5, axis=0),
    })

births_df = summarize_component(births_tot)
deaths_df = summarize_component(deaths_tot)
netmig_df = summarize_component(mig_tot)
natinc_df = summarize_component(nat_inc)

def pick_years(df, years=(2024, 2030, 2035, 2040, 2045), scale=1000.0):
    """
    Выбирает ключевые годы прогноза и переводит показатели в тыс. человек.
    """
    out = df[df["Year"].isin(years)].copy()

    for c in ["Median", "P5", "P95", "P12_5", "P87_5"]:
        out[c] = (out[c] / scale).round(2)

```

```

return out

print("Число рождений, тыс. человек:")
display(pick_years(births_df))

print("Число умерших, тыс. человек:")
display(pick_years(deaths_df))

print("Чистая миграция, применённая в прогнозной модели, тыс. человек:")
display(pick_years(netmig_df))

print("Естественный прирост населения, тыс. человек:")
display(pick_years(natinc_df))

```

```

# === Диагностика ARIMA-спецификаций для рядов главных компонент ===

import warnings
warnings.filterwarnings("ignore")

import numpy as np
import pandas as pd
from statsmodels.tsa.arima.model import ARIMA

# Спецификации ARIMA, включённые в сравнение.
CANDIDATES = [
    (1,0,0),
    (2,0,0),
    (1,0,1),
    (0,1,1),
    (1,1,0),
    (1,1,1),
]

# Базовая спецификация, используемая в основной модели.
BASE_ORDER = (1,0,0)

def fit_aic_bic(series, order):
    """
    Оценивает ARIMA-модель заданной спецификации
    и возвращает значения AIC и BIC.
    """
    m = ARIMA(series, order=order)
    r = m.fit()
    return float(r.aic), float(r.bic)

def evaluate_model_block(block_name, model_dict_or_obj, sexes=None):
    """
    Сравнивает ARIMA-спецификации для рядов главных компонент.

    model_dict_or_obj:
    - словарь по полу для моделей смертности и миграции;
    - отдельный объект модели для рождаемости.
    """
    rows = []

```

```

if sexes is None:
    items = [("NA", model_dict_or_obj)]
else:
    items = [(sex, model_dict_or_obj[sex]) for sex in sexes]

for sex, mdl in items:
    scores = mdl["scores"]
    n_comp = scores.shape[1]

    for pc in range(n_comp):
        series = np.asarray(scores[:, pc], dtype=float)

        # Проверка корректности ряда главной компоненты.
        if not np.isfinite(series).all():
            rows.append({
                "block": block_name,
                "sex": sex,
                "pc": pc + 1,
                "best_aic_order": None,
                "best_aic": np.nan,
                "best_bic_order": None,
                "best_bic": np.nan,
                "aic_base": np.nan,
                "bic_base": np.nan,
                "delta_aic": np.nan,
                "delta_bic": np.nan,
                "note": "non-finite values in series"
            })
            continue

        # Оценка базовой спецификации ARIMA(1,0,0).
        try:
            aic_base, bic_base = fit_aic_bic(series, BASE_ORDER)
        except Exception:
            aic_base, bic_base = np.nan, np.nan

        best_aic = (None, np.inf)
        best_bic = (None, np.inf)

        # Перебор альтернативных спецификаций.
        for order in CANDIDATES:
            try:
                aic, bic = fit_aic_bic(series, order)

                if aic < best_aic[1]:
                    best_aic = (order, aic)

                if bic < best_bic[1]:
                    best_bic = (order, bic)

            except Exception:
                continue

        rows.append({
            "block": block_name,
            "sex": sex,
            "pc": pc + 1,
            "best_aic_order": best_aic[0],
            "best_aic": best_aic[1] if best_aic[0] is not None else np.nan,

```

```

        "best_bic_order": best_bic[0],
        "best_bic": best_bic[1] if best_bic[0] is not None else np.nan,
        "aic_base": aic_base,
        "bic_base": bic_base,
        "delta_aic": (
            best_aic[1] - aic_base
            if best_aic[0] is not None and np.isfinite(aic_base)
            else np.nan
        ),
        "delta_bic": (
            best_bic[1] - bic_base
            if best_bic[0] is not None and np.isfinite(bic_base)
            else np.nan
        ),
        "note": ""
    })

    return rows

all_rows = []
all_rows += evaluate_model_block("Mortality", mort_models, sexes=["F", "M"])
all_rows += evaluate_model_block("Migration", mig_models, sexes=["F", "M"])
all_rows += evaluate_model_block("Fertility", fert_model, sexes=None)

df = pd.DataFrame(all_rows)

# Сортировка результатов диагностики.
df_sorted = df.sort_values(["block", "sex", "pc"])

print("=== Сравнение ARIMA-спецификаций по AIC и BIC ===")
print("delta_aic и delta_bic рассчитаны относительно ARIMA(1,0,0); отрицательное значение означает улучшение относительно базовой спецификации.")
display(df_sorted)

print("\n=== Сводка по порядку интегрирования d для лучшей спецификации по AIC ===")

def get_d(order):
    return None if order is None else order[1]

summary = (
    df_sorted.assign(best_aic_d=df_sorted["best_aic_order"].apply(get_d))
    .groupby(["block", "sex"])["best_aic_d"]
    .value_counts(dropna=False)
    .rename("count")
    .reset_index()
)

display(summary)

df_sorted.to_excel("arima_specification_diagnostics.xlsx", index=False)
summary.to_excel("arima_specification_diagnostics_summary.xlsx", index=False)

print("Сохранены файлы:")
print("arima_specification_diagnostics.xlsx")
print("arima_specification_diagnostics_summary.xlsx")

```

```

# === Диагностика ARIMA-спецификаций для рядов главных компонент за 2000–2023
# ГГ. ===

import warnings
warnings.filterwarnings("ignore")

import numpy as np
import pandas as pd
from statsmodels.tsa.arima.model import ARIMA

# Спецификации ARIMA, включённые в сравнение.
CANDIDATES = [
    (1, 0, 0),
    (2, 0, 0),
    (1, 0, 1),
    (0, 1, 1),
    (1, 1, 0),
    (1, 1, 1),
]

# Базовая спецификация, используемая в основной модели.
BASE_ORDER = (1, 0, 0)

# Годы, соответствующие строкам матриц scores.
years = np.arange(1990, 2024)

# Период диагностики.
DIAG_START, DIAG_END = 2000, 2023
mask = (years >= DIAG_START) & (years <= DIAG_END)

def fit_aic_bic(series, order):
    """
    Оценивает ARIMA-модель заданной спецификации
    и возвращает значения AIC и BIC.
    """
    m = ARIMA(series, order=order)
    r = m.fit()
    return float(r.aic), float(r.bic)

def evaluate_model_block(block_name, model_dict_or_obj, sexes=None):
    """
    Сравнивает ARIMA-спецификации для рядов главных компонент
    за выбранный диагностический период.

    model_dict_or_obj:
    - словарь по полу для моделей смертности и миграции;
    - отдельный объект модели для рождаемости.
    """
    rows = []

    if sexes is None:
        items = [("NA", model_dict_or_obj)]
    else:
        items = [(sex, model_dict_or_obj[sex]) for sex in sexes]

```

```

for sex, mdl in items:
    scores = mdl["scores"]
    n_comp = scores.shape[1]

    # Обрезка рядов главных компонент по диагностическому периоду.
    scores_sub = scores[mask, :]

    for pc in range(n_comp):
        series = np.asarray(scores_sub[:, pc], dtype=float)

        # Проверка корректности ряда главной компоненты.
        if not np.isfinite(series).all():
            rows.append({
                "block": block_name,
                "sex": sex,
                "pc": pc + 1,
                "best_aic_order": None,
                "best_aic": np.nan,
                "best_bic_order": None,
                "best_bic": np.nan,
                "aic_base": np.nan,
                "bic_base": np.nan,
                "delta_aic": np.nan,
                "delta_bic": np.nan,
                "note": "non-finite values in series"
            })
            continue

        # Оценка базовой спецификации ARIMA(1,0,0).
        try:
            aic_base, bic_base = fit_aic_bic(series, BASE_ORDER)
        except Exception:
            aic_base, bic_base = np.nan, np.nan

        best_aic = (None, np.inf)
        best_bic = (None, np.inf)

        # Перебор альтернативных спецификаций.
        for order in CANDIDATES:
            try:
                aic, bic = fit_aic_bic(series, order)

                if aic < best_aic[1]:
                    best_aic = (order, aic)

                if bic < best_bic[1]:
                    best_bic = (order, bic)

            except Exception:
                continue

        rows.append({
            "block": block_name,
            "sex": sex,
            "pc": pc + 1,
            "best_aic_order": best_aic[0],
            "best_aic": best_aic[1] if best_aic[0] is not None else np.nan,
            "best_bic_order": best_bic[0],
            "best_bic": best_bic[1] if best_bic[0] is not None else np.nan,

```

```

        "aic_base": aic_base,
        "bic_base": bic_base,
        "delta_aic": (
            best_aic[1] - aic_base
            if best_aic[0] is not None and np.isfinite(aic_base)
            else np.nan
        ),
        "delta_bic": (
            best_bic[1] - bic_base
            if best_bic[0] is not None and np.isfinite(bic_base)
            else np.nan
        ),
        "note": ""
    })

    return rows

all_rows = []
all_rows += evaluate_model_block("Mortality", mort_models, sexes=["F", "M"])
all_rows += evaluate_model_block("Migration", mig_models, sexes=["F", "M"])
all_rows += evaluate_model_block("Fertility", fert_model, sexes=None)

df = pd.DataFrame(all_rows).sort_values(["block", "sex", "pc"])

print(f"=== {DIAG_START}-{DIAG_END}: сравнение ARIMA-спецификаций по AIC и BIC ===")
print("delta_aic и delta_bic рассчитаны относительно ARIMA(1,0,0); отрицательное значение означает улучшение относительно базовой спецификации.")
display(df)

print(f"\n=== {DIAG_START}-{DIAG_END}: сводка по порядку интегрирования d для лучшей спецификации по AIC ===")

def get_d(order):
    return None if order is None else order[1]

summary = (
    df.assign(best_aic_d=df["best_aic_order"].apply(get_d))
    .groupby(["block", "sex"])["best_aic_d"]
    .value_counts(dropna=False)
    .rename("count")
    .reset_index()
)

display(summary)

```

```

# === Проверка устойчивости результатов: период калибровки 2000-2023 гг. ===
# Диагностический блок не изменяет основную модель и не влияет на итоговый прогноз,
# если выполняется после основных расчётов.

```

```

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import warnings

```

```
warnings.filterwarnings("ignore")

ROBUST_START = 2000
ROBUST_DRAWS = N_SIMS

print(f"\n=== Проверка устойчивости: калибровка {ROBUST_START}-2023 гг. ===")

# -----
# 1. Ограничение исторического периода
# -----

mort_w_sub = mort_w.loc[mort_w.index >= ROBUST_START].copy()
mig_net_w_sub = mig_net_w.loc[mig_net_w.index >= ROBUST_START].copy()
fert_w_sub = fert_w.loc[fert_w.index >= ROBUST_START].copy()

# -----
# 2. Повторная оценка PCA + ARIMA на укороченном периоде
# -----

mort_models_sub = {}
mig_models_sub = {}

for sex in ["F", "M"]:
    mort_sub = mort_w_sub.xs(sex, axis=1, level=0).copy()
    mort_models_sub[sex] = pca_arima_simulation(
        mort_sub,
        n_comp=2,
        draws=ROBUST_DRAWS
    )

    mig_sub = mig_net_w_sub.xs(sex, axis=1, level=0).copy()
    mig_models_sub[sex] = pca_arima_simulation(
        mig_sub,
        n_comp=2,
        draws=ROBUST_DRAWS
    )

fert_model_sub = pca_arima_simulation(
    fert_w_sub,
    n_comp=2,
    draws=ROBUST_DRAWS
)

# -----
# 3. Вспомогательные функции
# -----

def summarize_scalar_sims(sim_NT, years):
    """
    sim_NT: массив  $N \times T$ .
    Возвращает таблицу Year, Median, P10 и P90.
    """
    return pd.DataFrame({
        "Year": years,
        "Median": np.median(sim_NT, axis=0),
        "P10": np.percentile(sim_NT, 10, axis=0),
        "P90": np.percentile(sim_NT, 90, axis=0),
    })
```

```

def summarize_mortality_mean(model_obj, last_hist_year):
    """
    Для смертности рассчитывается средний возрастной коэффициент
    по всем возрастным группам.
    """
    sims = np.asarray(model_obj["sims"], dtype=float)
    years_fore = np.arange(last_hist_year + 1, last_hist_year + 1 +
sims.shape[1])
    sim_mean = sims.mean(axis=2)

    return summarize_scalar_sims(sim_mean, years_fore)

def summarize_fertility_tfr(model_obj, last_hist_year):
    """
    Для рождаемости рассчитывается суммарный коэффициент рождаемости
    на основе возрастных коэффициентов рождаемости.
    """
    sims = np.asarray(model_obj["sims"], dtype=float)
    years_fore = np.arange(last_hist_year + 1, last_hist_year + 1 +
sims.shape[1])

    asfr_labels = [str(c) for c in fert_w.columns]
    repro_ok = [g for g in REPRO_AGE_GROUPS if g in asfr_labels]

    if len(repro_ok) < 5:
        raise ValueError(f"Недостаточно репродуктивных групп для расчёта TFR:
{repro_ok}")

    idx = np.array([asfr_labels.index(g) for g in repro_ok], dtype=int)

    tfr_sims = sims[:, :, idx].sum(axis=2) * 5.0 / 1000.0

    return summarize_scalar_sims(tfr_sims, years_fore)

def plot_compare(df_full, df_sub, title, ylabel, label_full, label_sub,
filename=None):
    plt.figure(figsize=(10, 5))

    # Прогноз по полному периоду калибровки.
    plt.plot(
        df_full["Year"],
        df_full["Median"],
        linewidth=2.2,
        color="#000000",
        label=label_full
    )

    plt.fill_between(
        df_full["Year"],
        df_full["P10"],
        df_full["P90"],
        alpha=1.0,
        color="#E6E6E6"
    )

```

```

# Прогноз по укороченному периоду калибровки.
plt.plot(
    df_sub["Year"],
    df_sub["Median"],
    linewidth=2.2,
    color="#6F6F6F",
    linestyle="--",
    label=label_sub
)

plt.fill_between(
    df_sub["Year"],
    df_sub["P10"],
    df_sub["P90"],
    alpha=1.0,
    color="#C9C9C9"
)

plt.title(title)
plt.xlabel("Год")
plt.ylabel(ylabel)
plt.grid(alpha=0.3)
plt.legend()
plt.tight_layout()

if filename is not None:
    plt.savefig(filename, dpi=300, bbox_inches="tight")

plt.show()

# -----
# 4. Формирование сводок по полному и укороченному периодам
# -----

mortF_full_df = summarize_mortality_mean(
    mort_models["F"],
    last_hist_year=int(mort_w.index.max())
)

mortF_sub_df = summarize_mortality_mean(
    mort_models_sub["F"],
    last_hist_year=int(mort_w_sub.index.max())
)

fert_full_df = summarize_fertility_tfr(
    fert_model,
    last_hist_year=int(fert_w.index.max())
)

fert_sub_df = summarize_fertility_tfr(
    fert_model_sub,
    last_hist_year=int(fert_w_sub.index.max())
)

# -----
# 5. Сравнение последнего года прогнозного горизонта
# -----

print("\n--- Сравнение последнего года прогноза ---")

```

```

print(
    "Смертность женщин, средний возрастной коэффициент, 2045:\n"
    f" 1990-2023: {mortF_full_df.loc[mortF_full_df['Year'] ==
mortF_full_df['Year'].max(), 'Median'].iloc[0]:.3f}\n"
    f" 2000-2023: {mortF_sub_df.loc[mortF_sub_df['Year'] ==
mortF_sub_df['Year'].max(), 'Median'].iloc[0]:.3f}"
)

print(
    "TFR, 2045:\n"
    f" 1990-2023: {fert_full_df.loc[fert_full_df['Year'] ==
fert_full_df['Year'].max(), 'Median'].iloc[0]:.3f}\n"
    f" 2000-2023: {fert_sub_df.loc[fert_sub_df['Year'] ==
fert_sub_df['Year'].max(), 'Median'].iloc[0]:.3f}"
)

# -----
# 6. Построение графиков
# -----

plot_compare(
    df_full=mortF_full_df,
    df_sub=mortF_sub_df,
    title="Смертность женщин: сравнение прогнозов при разных периодах
калибровки",
    ylabel="Средний возрастной коэффициент смертности, %",
    label_full="Калибровка 1990-2023",
    label_sub="Калибровка 2000-2023",
    filename="robustness_mortality_F_2000_2023.png"
)

plot_compare(
    df_full=fert_full_df,
    df_sub=fert_sub_df,
    title="СКР: сравнение прогнозов при разных периодах калибровки",
    ylabel="Детей на женщину",
    label_full="Калибровка 1990-2023",
    label_sub="Калибровка 2000-2023",
    filename="robustness_tfr_2000_2023.png"
)

# -----
# 7. Сохранение таблиц
# -----

with pd.ExcelWriter("robustness_check_2000_2023.xlsx", engine="openpyxl") as
writer:
    mortF_full_df.to_excel(writer, sheet_name="Mortality_F_1990_2023",
index=False)
    mortF_sub_df.to_excel(writer, sheet_name="Mortality_F_2000_2023",
index=False)
    fert_full_df.to_excel(writer, sheet_name="TFR_1990_2023", index=False)
    fert_sub_df.to_excel(writer, sheet_name="TFR_2000_2023", index=False)

print("\nСохранены файлы:")
print("robustness_mortality_F_2000_2023.png")
print("robustness_tfr_2000_2023.png")
print("robustness_check_2000_2023.xlsx")

```

```

# =====
# Коэффициенты демографической нагрузки
# возрастные группы: 0-14, 15-64, 65 лет и старше
# =====

import numpy as np
import pandas as pd

# -----
# 1. Определение возрастных групп
# -----

def age_start(age_group):
    """
    Возвращает нижнюю границу возрастной группы.
    Например: '0-4' -> 0, '65-69' -> 65, '85+' -> 85.
    """
    age_group = str(age_group).strip()

    if age_group.endswith("+"):
        return int(age_group[:-1])

    return int(age_group.split("-")[0])

age_starts = np.array([age_start(a) for a in AGE_LABELS])

idx_child = np.where(age_starts < 15)[0]
idx_work  = np.where((age_starts >= 15) & (age_starts < 65))[0]
idx_old   = np.where(age_starts >= 65)[0]

# -----
# 2. Расчёт численности по укрупнённым возрастным группам
# -----

pop_sim_tot = pop_sim_F + pop_sim_M # N × T × A

N, T, A = pop_sim_tot.shape

try:
    YEARS
except NameError:
    YEARS = np.arange(2024, 2024 + T)

child = pop_sim_tot[:, :, idx_child].sum(axis=2)
work  = pop_sim_tot[:, :, idx_work].sum(axis=2)
old   = pop_sim_tot[:, :, idx_old].sum(axis=2)

# Защита от деления на ноль.
work = np.clip(work, 1e-9, None)

# -----
# 3. Расчёт коэффициентов демографической нагрузки
# -----

dr_total = (child + old) / work * 100.0
dr_child = child / work * 100.0

```

```

dr_old    = old / work * 100.0

# -----
# 4. Формирование итоговой таблицы
# -----

dep_df = pd.DataFrame({
    "Year": YEARS,

    "DR_total_median": np.median(dr_total, axis=0),
    "DR_total_P10": np.percentile(dr_total, 10, axis=0),
    "DR_total_P90": np.percentile(dr_total, 90, axis=0),

    "DR_child_median": np.median(dr_child, axis=0),
    "DR_child_P10": np.percentile(dr_child, 10, axis=0),
    "DR_child_P90": np.percentile(dr_child, 90, axis=0),

    "DR_old_median": np.median(dr_old, axis=0),
    "DR_old_P10": np.percentile(dr_old, 10, axis=0),
    "DR_old_P90": np.percentile(dr_old, 90, axis=0),
})

dep_df.to_excel("dependency_ratio.xlsx", index=False)

print("Файл сохранён: dependency_ratio.xlsx")
display(dep_df.head())

```

```

# =====
# Компоненты базовой вероятностной модели:
# рождаемость, смертность, чистая миграция, СКР и общие коэффициенты
# =====

import numpy as np
import pandas as pd

# -----
# 1. Вспомогательные функции
# -----

def safe_vec(x, n, name):
    """
    Проверяет, что вектор имеет требуемую длину.
    """
    v = np.asarray(x, dtype=float).reshape(-1)

    if v.size != n:
        raise ValueError(f"{name}: размер {v.size}, ожидалось {n}")

    return v

def tfr_from_asfr_promille(asfr_vec):
    """
    Рассчитывает суммарный коэффициент рождаемости на основе ASFR.
    ASFR заданы в промилле, возрастные группы являются пятилетними.
    """
    s = pd.Series(
        np.asarray(asfr_vec, float),

```

```

        index=[str(x) for x in ASFR_LABELS]
    ).fillna(0.0)

    return float(
        s.reindex(REPRO_AGE_GROUPS).fillna(0.0).sum() * 5.0 / 1000.0
    )

def q(x):
    """
    Возвращает медиану, P10 и P90.
    """
    return np.median(x), np.percentile(x, 10), np.percentile(x, 90)

def build_annual_transfer_safe(n_age):
    """
    Формирует матрицу годового перехода между пятилетними возрастными группами:
    4/5 численности остаётся в возрастной группе,
    1/5 переходит в следующую возрастную группу.
    Старшая возрастная группа является открытой.
    """
    Tm = np.zeros((n_age, n_age))

    for i in range(n_age):
        stay = 1.0 if i == n_age - 1 else 4.0 / 5.0
        Tm[i, i] += stay

        if i < n_age - 1:
            Tm[i + 1, i] += 1.0 / 5.0

    return Tm

# -----
# 2. Размерности и стартовая возрастно-половая структура
# -----

N = int(N_SIMS)
T = len(YEARS)
A_LEN = len(AGE_LABELS)

p_male = SRB / (1 + SRB)
p_fem = 1 - p_male

base_F = safe_vec(
    pop_wide["F"].astype(float).values,
    A_LEN,
    "base_F"
)

base_M = safe_vec(
    pop_wide["M"].astype(float).values,
    A_LEN,
    "base_M"
)

# Матрица годового перехода формируется отдельно,
# чтобы не зависеть от возможного переопределения переменной A.

```

```
AGE_TRANSFER = build_annual_transfer_safe(A_LEN)

print("Размерность AGE_TRANSFER:", AGE_TRANSFER.shape)
print("Число возрастных групп:", A_LEN)
```

```
# -----
# 3. Массивы для хранения компонент
# -----
```

```
births_all = np.zeros((N, T))
deaths_all = np.zeros((N, T))
netmig_all = np.zeros((N, T))
tfr_all = np.zeros((N, T))
pop_end_all = np.zeros((N, T))
```

```
# -----
# 4. Основной цикл по симуляциям и годам
# -----
```

```
for d in range(N):
    popF = base_F.copy()
    popM = base_M.copy()

    for t, year in enumerate(YEARS):

        mortF = safe_vec(
            mort_models["F"]["sims"][d, t, :],
            A_LEN,
            "mortF"
        )

        mortM = safe_vec(
            mort_models["M"]["sims"][d, t, :],
            A_LEN,
            "mortM"
        )

        migF = safe_vec(
            mig_models["F"]["sims"][d, t, :],
            A_LEN,
            "migF"
        )

        migM = safe_vec(
            mig_models["M"]["sims"][d, t, :],
            A_LEN,
            "migM"
        )

        asfr = fert_model["sims"][d, t, :]

        # Расчёт СКР.
        tfr_all[d, t] = tfr_from_asfr_promille(asfr)

        # Расчёт смертности и годового перехода.
        survF = safe_vec(
            surv_prob_exp_by_age(mortF),
```

```

        A_LEN,
        "survF"
    )

    survM = safe_vec(
        surv_prob_exp_by_age(mortM),
        A_LEN,
        "survM"
    )

    deathsF = float(np.sum(popF * (1.0 - survF)))
    deathsM = float(np.sum(popM * (1.0 - survM)))
    deaths_all[d, t] = deathsF + deathsM

    popF = AGE_TRANSFER @ (popF * survF)
    popM = AGE_TRANSFER @ (popM * survM)

    # Расчёт чистой миграции, применённой в модели.
    migF_sh = shape_migration_sex(
        migF,
        "F",
        cap=MIG_GROUP_CAP
    )

    migM_sh = shape_migration_sex(
        migM,
        "M",
        cap=MIG_GROUP_CAP
    )

    bumpF = (
        MIG_NET_BUMP *
        MIG_BUMP_SEX_SPLIT[0]
    ) * bump_weights.values

    bumpM = (
        MIG_NET_BUMP *
        MIG_BUMP_SEX_SPLIT[1]
    ) * bump_weights.values

    popF = popF + migF_sh + bumpF
    popM = popM + migM_sh + bumpM

    netmig_all[d, t] = float(
        np.sum(migF_sh) +
        np.sum(migM_sh) +
        np.sum(bumpF) +
        np.sum(bumpM)
    )

    # Расчёт числа рождений.
    B = float(births_from_asfr(popF, asfr))
    births_all[d, t] = B

    popF[0] += B * p_fem
    popM[0] += B * p_male

    # Исключение отрицательных значений численности.
    popF = np.maximum(popF, 0)

```

```

popM = np.maximum(popM, 0)

# Численность населения на конец года.
pop_end_all[d, t] = float(np.sum(popF) + np.sum(popM))

# -----
# 5. Формирование сводной таблицы по годам
# -----

rows = []

for i, y in enumerate(YEARS):
    b_med, b_p10, b_p90 = q(births_all[:, i])
    d_med, d_p10, d_p90 = q(deaths_all[:, i])
    m_med, m_p10, m_p90 = q(netmig_all[:, i])
    t_med, t_p10, t_p90 = q(tfr_all[:, i])
    p_med, p_p10, p_p90 = q(pop_end_all[:, i])

    rows.append({
        "Year": y,

        "TFR_Median": t_med,
        "TFR_P10": t_p10,
        "TFR_P90": t_p90,

        "Births_Median": b_med,
        "Births_P10": b_p10,
        "Births_P90": b_p90,

        "Deaths_Median": d_med,
        "Deaths_P10": d_p10,
        "Deaths_P90": d_p90,

        "NetMig_Median": m_med,
        "NetMig_P10": m_p10,
        "NetMig_P90": m_p90,

        "PopEnd_Median": p_med,
        "PopEnd_P10": p_p10,
        "PopEnd_P90": p_p90,

        "CBR_per1000": b_med / p_med * 1000.0,
        "CDR_per1000": d_med / p_med * 1000.0,
        "NMR_per1000": m_med / p_med * 1000.0,

        "Natural_Increase": b_med - d_med,
        "Total_Balance": (b_med - d_med) + m_med,
    })

comp_dyn_OK = pd.DataFrame(rows)

comp_dyn_OK.to_excel("component_dynamics_OK.xlsx", index=False)

print("Файл сохранён: component_dynamics_OK.xlsx")
display(comp_dyn_OK.head())
display(comp_dyn_OK.tail())

```

```

# =====
# Расчёт ожидаемой продолжительности жизни на основе
# модельных возрастных коэффициентов смертности
# =====

import numpy as np
import pandas as pd

# -----
# 1. Параметры краткой таблицы смертности
# -----

def age_width(age_label):
    """
    Возвращает ширину возрастного интервала.
    Для открытой старшей возрастной группы возвращается np.inf.
    """
    age_label = str(age_label)

    if age_label.endswith("+"):
        return np.inf

    a, b = age_label.split("-")
    return int(b) - int(a) + 1

WIDTHS = np.array([age_width(a) for a in AGE_LABELS], dtype=float)

# Нижняя граница возрастной группы.
AGE_LO = np.array([
    int(str(a).split("-")[0].replace("+", ""))
    for a in AGE_LABELS
])

# Среднее число лет, прожитых умершими внутри возрастного интервала.
# Для пятилетних групп используется приближение n/2.
AX = np.where(np.isfinite(WIDTHS), WIDTHS / 2, np.nan)

# Для группы 0-4 сохраняется нейтральное приближение,
# поскольку младенческая смертность отдельно не выделяется.
AX[0] = 2.5

def prepare_mortality_rates(m_promille, apply_model_adjustments=True):
    """
    Преобразует возрастные коэффициенты смертности из промилле
    в доли на одного человека.
    """
    m = np.asarray(m_promille, dtype=float).copy()

    if apply_model_adjustments:
        # Применяется та же логика корректировки, что и в прогнозной модели:
        # для группы 85+ используется множитель OLD_AGE_MORT_MULT,
        # затем применяются возрастные ограничения CAP_BY_AGE.
        m = m * np.where(AGE_LO >= 85, OLD_AGE_MORT_MULT, 1.0)
        m = np.clip(m, 0.0, CAP_BY_AGE)

    return m / 1000.0

```

```

def abridged_life_expectancy_from_mx(m_promille, radix=100000,
    apply_model_adjustments=True):
    """
    Рассчитывает ожидаемую продолжительность жизни при рождении
    по краткой таблице смертности на основе пятилетних возрастных коэффициентов.
    """
    mx = prepare_mortality_rates(
        m_promille,
        apply_model_adjustments=apply_model_adjustments
    )

    n = WIDTHS.copy()
    ax = AX.copy()

    qx = np.zeros_like(mx, dtype=float)

    # Закрытые возрастные интервалы.
    closed = np.isfinite(n)
    qx[closed] = (
        n[closed] * mx[closed]
    ) / (
        1 + (n[closed] - ax[closed]) * mx[closed]
    )
    qx[closed] = np.clip(qx[closed], 0, 1)

    # Открытая старшая возрастная группа.
    qx[~closed] = 1.0

    lx = np.zeros_like(mx, dtype=float)
    dx = np.zeros_like(mx, dtype=float)
    Lx = np.zeros_like(mx, dtype=float)

    lx[0] = radix

    for i in range(len(mx)):
        dx[i] = lx[i] * qx[i]

        if np.isfinite(n[i]):
            Lx[i] = n[i] * lx[i] - (n[i] - ax[i]) * dx[i]
        else:
            # Для открытого интервала используется приближение  $L\omega = l\omega / m\omega$ .
            Lx[i] = lx[i] / max(mx[i], 1e-9)

        if i < len(mx) - 1:
            lx[i + 1] = lx[i] - dx[i]

    Tx = np.cumsum(Lx[::-1])[::-1]
    e0 = Tx[0] / radix

    return e0

# =====
# 2. Историческая оценка ОПЖ за 2023 год
# =====

e0_hist = []

```

```

for sex in ["M", "F"]:
    m_2023 = mort_w.xs(sex, axis=1, level=0).loc[2023].values

    e0 = abridged_life_expectancy_from_mx(
        m_2023,
        apply_model_adjustments=True
    )

    e0_hist.append({
        "Year": 2023,
        "Sex": sex,
        "e0": e0,
        "Type": "historical_from_model_mx"
    })

e0_hist_df = pd.DataFrame(e0_hist)

# =====
# 3. Прогнозная ОПЖ по стохастическим траекториям смертности
# =====

def e0_distribution_for_year(sex, year):
    """
    Рассчитывает распределение e0 по всем стохастическим траекториям
    для заданного пола и года прогноза.
    """
    t = YEARS.index(year)
    sims = mort_models[sex]["sims"][:, t, :]

    e0_values = np.array([
        abridged_life_expectancy_from_mx(
            row,
            apply_model_adjustments=True
        )
        for row in sims
    ])

    return e0_values

rows = []

for year in [2024, 2030, 2035, 2040, 2045]:
    for sex in ["M", "F"]:
        vals = e0_distribution_for_year(sex, year)

        rows.append({
            "Year": year,
            "Sex": sex,
            "e0_median": np.median(vals),
            "e0_p10": np.percentile(vals, 10),
            "e0_p90": np.percentile(vals, 90),
            "e0_mean": np.mean(vals)
        })

e0_forecast_df = pd.DataFrame(rows)

```

```

# =====
# 4. Таблица медианной ОПЖ по полу и разрыва между женщинами и мужчинами
# =====

e0_wide = (
    e0_forecast_df
        .pivot(index="Year", columns="Sex", values="e0_median")
        .reset_index()
)

e0_wide["Gap_F_minus_M"] = e0_wide["F"] - e0_wide["M"]

print("Историческая оценка ОПЖ за 2023 год на основе модельных mx:")
display(e0_hist_df)

print("Прогнозная ОПЖ: медиана и интервал P10-P90:")
display(e0_forecast_df)

print("Медианная ОПЖ и разрыв между женщинами и мужчинами:")
display(e0_wide)

# =====
# 5. Сохранение результатов
# =====

e0_hist_df.to_excel("life_expectancy_2023_from_model_mx.xlsx", index=False)
e0_forecast_df.to_excel("life_expectancy_forecast_from_mortality_sims.xlsx",
index=False)
e0_wide.to_excel("life_expectancy_forecast_wide_gap.xlsx", index=False)

print("Сохранены файлы:")
print("life_expectancy_2023_from_model_mx.xlsx")
print("life_expectancy_forecast_from_mortality_sims.xlsx")
print("life_expectancy_forecast_wide_gap.xlsx")

```

```

# =====
# Проверка чувствительности расчёта ОПЖ
# к условному разделению группы 0-4 на 0 и 1-4
# =====

import numpy as np
import pandas as pd

# -----
# 1. Параметры расчёта
# -----

# Младенческая смертность, промилле.
INFANT_MORTALITY_PROMILLE = 4.1

# Среднее число лет, прожитых умершими в возрасте 0 лет.
AX0 = 0.1

# -----
# 2. Формирование уточнённой возрастной сетки
# -----

```

```

AGE_LABELS_SPLIT = ["0", "1-4"] + AGE_LABELS[1:]

def age_width_split(age_label):
    """
    Возвращает ширину возрастного интервала
    для сетки 0, 1-4, 5-9, ..., 85+.
    """
    age_label = str(age_label)

    if age_label == "0":
        return 1

    if age_label == "1-4":
        return 4

    if age_label.endswith("+"):
        return np.inf

    a, b = age_label.split("-")
    return int(b) - int(a) + 1

WIDTHS_SPLIT = np.array(
    [age_width_split(a) for a in AGE_LABELS_SPLIT],
    dtype=float
)

AGE_LO_SPLIT = np.array([
    0 if str(a) == "0" else
    1 if str(a) == "1-4" else
    int(str(a).split("-")[0].replace("+", ""))
    for a in AGE_LABELS_SPLIT
])

AX_SPLIT = np.where(np.isfinite(WIDTHS_SPLIT), WIDTHS_SPLIT / 2, np.nan)
AX_SPLIT[0] = AX0
AX_SPLIT[1] = 2.0

# -----
# 3. Разделение смертности группы 0-4 на 0 и 1-4
# -----

def split_0_4_mortality(m_promille_5yr,
    infant_mortality_promille=INFANT_MORTALITY_PROMILLE):
    """
    Преобразует исходный вектор смертности по группам 0-4, 5-9, ...
    в вектор по группам 0, 1-4, 5-9, ...

    Для возраста 0 задаётся отдельное значение младенческой смертности.
    Для группы 1-4 уровень подбирается так, чтобы средний уровень
    смертности группы 0-4 приблизительно сохранялся.
    """
    m = np.asarray(m_promille_5yr, dtype=float).copy()

    m_0_4 = m[0]
    m_0 = infant_mortality_promille

```

```

# Сохранение среднего уровня смертности группы 0-4:
#  $5 * m_{0-4} \approx 1 * m_0 + 4 * m_{1-4}$ .
m_1_4 = (5 * m_0_4 - m_0) / 4
m_1_4 = max(m_1_4, 0.01)

m_split = np.concatenate([m_0, m_1_4], m[1:])

return m_split

# -----
# 4. Краткая таблица смертности для уточнённой сетки
# -----

def prepare_mortality_rates_split(m_promille_split,
    apply_model_adjustments=True):
    """
    Преобразует коэффициенты смертности по сетке 0, 1-4, 5-9, ..., 85+
    из промилле в доли на одного человека.
    """
    m = np.asarray(m_promille_split, dtype=float).copy()

    if apply_model_adjustments:
        # Для группы 85+ сохраняется та же корректировка, что и в основной
        # модели.
        m = m * np.where(AGE_LO_SPLIT >= 85, OLD_AGE_MORT_MULT, 1.0)

        # Возрастные ограничения расширяются на группы 0 и 1-4.
        cap_split = np.concatenate([[35.0, 35.0], CAP_BY_AGE[1:]])
        m = np.clip(m, 0.0, cap_split)

    return m / 1000.0

def abridged_life_expectancy_split(
    m_promille_5yr,
    infant_mortality_promille=INFANT_MORTALITY_PROMILLE,
    radix=100000,
    apply_model_adjustments=True
):
    """
    Рассчитывает ожидаемую продолжительность жизни при рождении
    после условного разделения группы 0-4 на 0 и 1-4.
    На вход подаётся исходный вектор смертности по группам 0-4, 5-9, ...
    """
    m_split = split_0_4_mortality(
        m_promille_5yr,
        infant_mortality_promille=infant_mortality_promille
    )

    mx = prepare_mortality_rates_split(
        m_split,
        apply_model_adjustments=apply_model_adjustments
    )

    n = WIDTHS_SPLIT.copy()
    ax = AX_SPLIT.copy()

    qx = np.zeros_like(mx, dtype=float)

```

```

closed = np.isfinite(n)

qx[closed] = (
    n[closed] * mx[closed]
) / (
    1 + (n[closed] - ax[closed]) * mx[closed]
)

qx[closed] = np.clip(qx[closed], 0, 1)

# Открытая старшая возрастная группа.
qx[~closed] = 1.0

lx = np.zeros_like(mx, dtype=float)
dx = np.zeros_like(mx, dtype=float)
Lx = np.zeros_like(mx, dtype=float)

lx[0] = radix

for i in range(len(mx)):
    dx[i] = lx[i] * qx[i]

    if np.isfinite(n[i]):
        Lx[i] = n[i] * lx[i] - (n[i] - ax[i]) * dx[i]
    else:
        Lx[i] = lx[i] / max(mx[i], 1e-9)

    if i < len(mx) - 1:
        lx[i + 1] = lx[i] - dx[i]

Tx = np.cumsum(Lx[::-1])[::-1]
e0 = Tx[0] / radix

return e0

# -----
# 5. Историческая оценка ОПЖ за 2023 год
# -----

rows_hist_split = []

for sex in ["M", "F"]:
    m_2023 = mort_w.xs(sex, axis=1, level=0).loc[2023].values

    e0_split = abridged_life_expectancy_split(
        m_2023,
        infant_mortality_promille=INFANT_MORTALITY_PROMILLE
    )

    rows_hist_split.append({
        "Year": 2023,
        "Sex": sex,
        "Infant_mortality_promille": INFANT_MORTALITY_PROMILLE,
        "e0_split_0_1_4": e0_split
    })

e0_hist_split_df = pd.DataFrame(rows_hist_split)

```

```

print("ОПЖ за 2023 год при условном разделении группы 0-4 на 0 и 1-4:")
display(e0_hist_split_df)

# -----
# 6. Прогнозная ОПЖ по стохастическим траекториям
# -----

def e0_distribution_for_year_split(sex, year,
infant_mortality_promille=INFANT_MORTALITY_PROMILLE):
    """
    Рассчитывает распределение ОПЖ по стохастическим траекториям
    для заданного пола и года прогноза.
    """
    t = YEARS.index(year)
    sims = mort_models[sex]["sims"][:, t, :]

    e0_values = np.array([
        abridged_life_expectancy_split(
            row,
            infant_mortality_promille=infant_mortality_promille
        )
        for row in sims
    ])

    return e0_values

rows_fore_split = []

for year in [2024, 2030, 2035, 2040, 2045]:
    for sex in ["M", "F"]:
        vals = e0_distribution_for_year_split(
            sex,
            year,
            infant_mortality_promille=INFANT_MORTALITY_PROMILLE
        )

        rows_fore_split.append({
            "Year": year,
            "Sex": sex,
            "Infant_mortality_promille": INFANT_MORTALITY_PROMILLE,
            "e0_median_split": np.median(vals),
            "e0_p10_split": np.percentile(vals, 10),
            "e0_p90_split": np.percentile(vals, 90),
            "e0_mean_split": np.mean(vals)
        })

e0_forecast_split_df = pd.DataFrame(rows_fore_split)

print("Прогнозная ОПЖ при условном разделении группы 0-4 на 0 и 1-4:")
display(e0_forecast_split_df)

# -----
# 7. Таблица медианной ОПЖ по полу и разрыва между женщинами и мужчинами
# -----

```

```

e0_split_wide = (
    e0_forecast_split_df
    .pivot(index="Year", columns="Sex", values="e0_median_split")
    .reset_index()
)

e0_split_wide["Gap_F_minus_M_split"] = e0_split_wide["F"] - e0_split_wide["M"]

print("Медианная ОПЖ и разрыв между женщинами и мужчинами:")
display(e0_split_wide)

# -----
# 8. Сохранение результатов
# -----

e0_hist_split_df.to_excel("life_expectancy_2023_split_0_1_4.xlsx", index=False)
e0_forecast_split_df.to_excel("life_expectancy_forecast_split_0_1_4.xlsx",
index=False)
e0_split_wide.to_excel("life_expectancy_forecast_split_0_1_4_wide_gap.xlsx",
index=False)

print("Сохранены файлы:")
print("life_expectancy_2023_split_0_1_4.xlsx")
print("life_expectancy_forecast_split_0_1_4.xlsx")
print("life_expectancy_forecast_split_0_1_4_wide_gap.xlsx")

```

```

# =====
# Сценарный анализ демографического развития Санкт-Петербурга
# =====

import numpy as np
import pandas as pd

# -----
# 1. Матрица годового перехода между возрастными группами
# -----

def build_annual_transfer_safe(n_age):
    """
    Формирует матрицу годового перехода между пятилетними возрастными группами:
    4/5 численности остаётся в возрастной группе,
    1/5 переходит в следующую возрастную группу.
    Старшая возрастная группа является открытой.
    """
    Tm = np.zeros((n_age, n_age))

    for i in range(n_age):
        stay = 1.0 if i == n_age - 1 else 4.0 / 5.0
        Tm[i, i] += stay

        if i < n_age - 1:
            Tm[i + 1, i] += 1.0 / 5.0

    return Tm

AGE_TRANSFER = build_annual_transfer_safe(len(AGE_LABELS))

```

```

# -----
# 2. Параметры сценариев
# -----

SCENARIOS_NEW = {
    "C1": {
        "label": "C1: благоприятный",
        "fert_mult": 1.05,
        "mort_improve_M": 0.012,
        "mort_improve_F": 0.009,
        "mig_bump": 35_000
    },
    "C2": {
        "label": "C2: умеренно благоприятный",
        "fert_mult": 1.03,
        "mort_improve_M": 0.008,
        "mort_improve_F": 0.006,
        "mig_bump": 25_000
    },
    "C3": {
        "label": "C3: умеренно неблагоприятный",
        "fert_mult": 0.98,
        "mort_improve_M": 0.003,
        "mort_improve_F": 0.002,
        "mig_bump": 15_000
    },
    "C4": {
        "label": "C4: неблагоприятный",
        "fert_mult": 0.95,
        "mort_improve_M": -0.003,
        "mort_improve_F": -0.002,
        "mig_bump": 0
    }
}

scenario_params_rows = []

for scen_name, params in SCENARIOS_NEW.items():
    scenario_params_rows.append({
        "Scenario": scen_name,
        "Scenario_label": params["label"],
        "ASFR_multiplier": params["fert_mult"],
        "ASFR_change_percent": (params["fert_mult"] - 1) * 100,
        "ASMR_annual_change_M_percent": -params["mort_improve_M"] * 100,
        "ASMR_annual_change_F_percent": -params["mort_improve_F"] * 100,
        "Migration_bump_persons": params["mig_bump"],
        "Migration_bump_thousand": params["mig_bump"] / 1000
    })

scenario_params_df = pd.DataFrame(scenario_params_rows)

# -----
# 3. Служебные параметры и вспомогательные функции
# -----

N_EXPORT = N_SIMS
YEARS_EXPORT = YEARS
T = len(YEARS_EXPORT)

```

```

A_LEN = len(AGE_LABELS)

p_male = SRB / (1 + SRB)
p_fem = 1 - p_male

base_F = pop_wide["F"].astype(float).values.copy()
base_M = pop_wide["M"].astype(float).values.copy()

AGE_LO = np.array([
    int(str(a).split("-")[0].replace("+", ""))
    for a in AGE_LABELS
])

def mortality_mult_by_year(year, sex, params):
    """
    Рассчитывает накопительное сценарное изменение возрастных коэффициентов
    смертности.
    Положительное значение mort_improve означает ежегодное снижение ASMR,
    отрицательное значение означает ежегодный рост ASMR.
    """
    h = year - 2023

    if sex == "M":
        r = params["mort_improve_M"]
    else:
        r = params["mort_improve_F"]

    return (1 - r) ** h

def surv_prob_exp_by_age_scenario(mort_rate_promille_vec, year, sex, params):
    """
    Рассчитывает вероятность дожития с учётом сценарного изменения смертности.
    """
    mult = mortality_mult_by_year(year, sex, params)
    m = np.asarray(mort_rate_promille_vec, float) * mult

    m = m * np.where(AGE_LO >= 85, OLD_AGE_MORT_MULT, 1.0)
    mx = np.clip(m, 0.0, CAP_BY_AGE) / 1000.0

    return np.exp(-mx)

def compute_tfr_from_asfr_vector(asfr_vec, asfr_labels):
    """
    Рассчитывает СКР на основе ASFR.
    ASFR заданы в промилле, возрастные группы являются пятилетними.
    """
    s = pd.Series(asfr_vec, index=asfr_labels, dtype=float)
    vals = s.reindex(REPRO_AGE_GROUPS).fillna(0.0)

    return vals.sum() * 5 / 1000

def age_width(age_label):
    """
    Возвращает ширину возрастного интервала.
    """

```

```

age_label = str(age_label)

if age_label.endswith("+"):
    return np.inf

a, b = age_label.split("-")
return int(b) - int(a) + 1

WIDTHS = np.array([age_width(a) for a in AGE_LABELS], dtype=float)
AX = np.where(np.isfinite(WIDTHS), WIDTHS / 2, np.nan)
AX[0] = 2.5

def abridged_life_expectancy_from_mx_scenario(
    m_promille,
    year,
    sex,
    params,
    radix=100000
):
    """
    Рассчитывает ожидаемую продолжительность жизни при рождении
    по краткой таблице смертности с учётом сценарного изменения ASMR.
    """
    mult = mortality_mult_by_year(year, sex, params)

    m = np.asarray(m_promille, dtype=float).copy() * mult
    m = m * np.where(AGE_LO >= 85, OLD_AGE_MORT_MULT, 1.0)
    m = np.clip(m, 0.0, CAP_BY_AGE)

    mx = m / 1000.0

    n = WIDTHS.copy()
    ax = AX.copy()

    qx = np.zeros_like(mx, dtype=float)

    closed = np.isfinite(n)

    qx[closed] = (
        n[closed] * mx[closed]
    ) / (
        1 + (n[closed] - ax[closed]) * mx[closed]
    )

    qx[closed] = np.clip(qx[closed], 0, 1)
    qx[~closed] = 1.0

    lx = np.zeros_like(mx, dtype=float)
    dx = np.zeros_like(mx, dtype=float)
    Lx = np.zeros_like(mx, dtype=float)

    lx[0] = radix

    for i in range(len(mx)):
        dx[i] = lx[i] * qx[i]

        if np.isfinite(n[i]):

```

```

        Lx[i] = n[i] * lx[i] - (n[i] - ax[i]) * dx[i]
    else:
        Lx[i] = lx[i] / max(mx[i], 1e-9)

    if i < len(mx) - 1:
        lx[i + 1] = lx[i] - dx[i]

    Tx = np.cumsum(Lx[::-1])[::-1]
    e0 = Tx[0] / radix

    return e0

# -----
# 4. Расчёт сценариев и сохранение компонент
# -----

population_rows = []
fertility_rows = []
mortality_deaths_rows = []
migration_rows = []
asmr_age_rows = []
e0_rows = []

YEARS_TO_CHECK = [2024, 2025, 2026, 2030, 2035, 2040, 2045]

for scen_name, params in SCENARIOS_NEW.items():

    print(f"Расчёт сценария {scen_name}")

    total_pop_sim = np.zeros((N_EXPORT, T))
    births_sim = np.zeros((N_EXPORT, T))
    deaths_sim = np.zeros((N_EXPORT, T))
    net_migration_sim = np.zeros((N_EXPORT, T))
    tfr_sim = np.zeros((N_EXPORT, T))

    asmr_M_adj = np.zeros((N_EXPORT, T, A_LEN))
    asmr_F_adj = np.zeros((N_EXPORT, T, A_LEN))

    for d in range(N_EXPORT):
        popF = base_F.copy()
        popM = base_M.copy()

        for t, year in enumerate(YEARS_EXPORT):

            mortF_raw = mort_models["F"]["sims"][d, t, :]
            mortM_raw = mort_models["M"]["sims"][d, t, :]

            migF_raw = mig_models["F"]["sims"][d, t, :]
            migM_raw = mig_models["M"]["sims"][d, t, :]

            asfr_raw = fert_model["sims"][d, t, :]
            asfr_adj = asfr_raw * params["fert_mult"]

            tfr_sim[d, t] = compute_tfr_from_asfr_vector(asfr_adj, ASFR_LABELS)

            multF = mortality_mult_by_year(year, "F", params)
            multM = mortality_mult_by_year(year, "M", params)

```

```

mortF_adj = mortF_raw * multF
mortM_adj = mortM_raw * multM

asmr_F_adj[d, t, :] = mortF_adj
asmr_M_adj[d, t, :] = mortM_adj

survF = surv_prob_exp_by_age_scenario(mortF_raw, year, "F", params)
survM = surv_prob_exp_by_age_scenario(mortM_raw, year, "M", params)

deathsF = np.sum(popF * (1 - survF))
deathsM = np.sum(popM * (1 - survM))
deaths_sim[d, t] = deathsF + deathsM

popF = AGE_TRANSFER @ (popF * survF)
popM = AGE_TRANSFER @ (popM * survM)

migF_shaped = shape_migration_sex(migF_raw, "F", cap=MIG_GROUP_CAP)
migM_shaped = shape_migration_sex(migM_raw, "M", cap=MIG_GROUP_CAP)

bumpF = (params["mig_bump"] * MIG_BUMP_SEX_SPLIT[0]) *
bump_weights.values
bumpM = (params["mig_bump"] * MIG_BUMP_SEX_SPLIT[1]) *
bump_weights.values

migF_total = migF_shaped + bumpF
migM_total = migM_shaped + bumpM

net_migration_sim[d, t] = np.sum(migF_total) + np.sum(migM_total)

popF += migF_total
popM += migM_total

B = births_from_asfr(popF, asfr_adj)
births_sim[d, t] = B

popF[0] += B * p_fem
popM[0] += B * p_male

popF = np.maximum(popF, 0)
popM = np.maximum(popM, 0)

total_pop_sim[d, t] = popF.sum() + popM.sum()

# Итоговые показатели по годам.
for t, year in enumerate(YEARS_EXPORT):

    population_rows.append({
        "Scenario": scen_name,
        "Scenario_label": params["label"],
        "Year": year,
        "Population_median": np.median(total_pop_sim[:, t]),
        "Population_p10": np.percentile(total_pop_sim[:, t], 10),
        "Population_p90": np.percentile(total_pop_sim[:, t], 90),
        "Population_p5": np.percentile(total_pop_sim[:, t], 5),
        "Population_p95": np.percentile(total_pop_sim[:, t], 95)
    })

    fertility_rows.append({
        "Scenario": scen_name,

```

```

        "Scenario_label": params["label"],
        "Year": year,
        "TFR_median": np.median(tfr_sim[:, t]),
        "TFR_p10": np.percentile(tfr_sim[:, t], 10),
        "TFR_p90": np.percentile(tfr_sim[:, t], 90),
        "Births_median": np.median(births_sim[:, t]),
        "Births_p10": np.percentile(births_sim[:, t], 10),
        "Births_p90": np.percentile(births_sim[:, t], 90)
    })

    mortality_deaths_rows.append({
        "Scenario": scen_name,
        "Scenario_label": params["label"],
        "Year": year,
        "Deaths_median": np.median(deaths_sim[:, t]),
        "Deaths_p10": np.percentile(deaths_sim[:, t], 10),
        "Deaths_p90": np.percentile(deaths_sim[:, t], 90)
    })

    migration_rows.append({
        "Scenario": scen_name,
        "Scenario_label": params["label"],
        "Year": year,
        "Net_migration_median": np.median(net_migration_sim[:, t]),
        "Net_migration_p10": np.percentile(net_migration_sim[:, t], 10),
        "Net_migration_p90": np.percentile(net_migration_sim[:, t], 90),
        "Migration_bump": params["mig_bump"]
    })

# Возрастно-половые ASMR.
for t, year in enumerate(YEARS_EXPORT):
    for i, age in enumerate(AGE_LABELS):

        asmr_age_rows.append({
            "Scenario": scen_name,
            "Scenario_label": params["label"],
            "Year": year,
            "Sex": "M",
            "Age_Group": age,
            "ASMR_median": np.median(asmr_M_adj[:, t, i]),
            "ASMR_p10": np.percentile(asmr_M_adj[:, t, i], 10),
            "ASMR_p90": np.percentile(asmr_M_adj[:, t, i], 90)
        })

        asmr_age_rows.append({
            "Scenario": scen_name,
            "Scenario_label": params["label"],
            "Year": year,
            "Sex": "F",
            "Age_Group": age,
            "ASMR_median": np.median(asmr_F_adj[:, t, i]),
            "ASMR_p10": np.percentile(asmr_F_adj[:, t, i], 10),
            "ASMR_p90": np.percentile(asmr_F_adj[:, t, i], 90)
        })

# ОПЖ по выбранным годам.
for year in YEARS_TO_CHECK:
    t = YEARS_EXPORT.index(year)

```

```

for sex in ["M", "F"]:
    sims = mort_models[sex]["sims"][:N_EXPORT, t, :]

    vals = np.array([
        abridged_life_expectancy_from_mx_scenario(
            row,
            year=year,
            sex=sex,
            params=params
        )
        for row in sims
    ])

    e0_rows.append({
        "Scenario": scen_name,
        "Scenario_label": params["label"],
        "Year": year,
        "Sex": sex,
        "Mortality_multiplier": mortality_mult_by_year(year, sex,
params),
        "e0_median": np.median(vals),
        "e0_p10": np.percentile(vals, 10),
        "e0_p90": np.percentile(vals, 90),
        "e0_mean": np.mean(vals)
    })

# -----
# 5. Формирование итоговых таблиц
# -----

scenario_population_export = pd.DataFrame(population_rows)
scenario_fertility_export = pd.DataFrame(fertility_rows)
scenario_mortality_deaths_export = pd.DataFrame(mortality_deaths_rows)
scenario_migration_export = pd.DataFrame(migration_rows)
scenario_asmr_age_export = pd.DataFrame(asmr_age_rows)
e0_scenario_new_long = pd.DataFrame(e0_rows)

e0_scenario_new_wide = (
    e0_scenario_new_long
    .pivot_table(
        index=["Scenario", "Scenario_label", "Year"],
        columns="Sex",
        values="e0_median"
    )
    .reset_index()
)

e0_scenario_new_wide["Gap_F_minus_M"] = (
    e0_scenario_new_wide["F"] - e0_scenario_new_wide["M"]
)

scenario_summary_all = (
    scenario_population_export
    .merge(
        scenario_fertility_export,
        on=["Scenario", "Scenario_label", "Year"],
        how="left"
    )
)

```

```

.merge(
    scenario_mortality_deaths_export,
    on=["Scenario", "Scenario_label", "Year"],
    how="left"
)
.merge(
    scenario_migration_export,
    on=["Scenario", "Scenario_label", "Year"],
    how="left"
)
.merge(
    e0_scenario_new_wide[
        ["Scenario", "Scenario_label", "Year", "M", "F", "Gap_F_minus_M"]
    ],
    on=["Scenario", "Scenario_label", "Year"],
    how="left"
)
)

scenario_summary_all = scenario_summary_all.rename(
    columns={
        "M": "M_e0",
        "F": "F_e0"
    }
)

summary_2045_all = scenario_summary_all.query("Year == 2045").copy()

# -----
# 6. Сохранение результатов
# -----

scenario_params_df.to_excel("scenario_parameters_FINAL.xlsx", index=False)
scenario_population_export.to_excel("scenario_population_total_FINAL.xlsx",
index=False)
scenario_fertility_export.to_excel("scenario_fertility_TFR_births_FINAL.xlsx",
index=False)
scenario_mortality_deaths_export.to_excel("scenario_mortality_deaths_FINAL.xlsx"
, index=False)
scenario_migration_export.to_excel("scenario_migration_net_FINAL.xlsx",
index=False)
scenario_asmr_age_export.to_excel("scenario_mortality_ASMR_by_age_sex_FINAL.xlsx
", index=False)
e0_scenario_new_long.to_excel("scenario_life_expectancy_new_mortality_trend_long
.xlsx", index=False)
e0_scenario_new_wide.to_excel("scenario_life_expectancy_new_mortality_trend_wide
.xlsx", index=False)
scenario_summary_all.to_excel("scenario_summary_all_indicators_FINAL.xlsx",
index=False)
summary_2045_all.to_excel("scenario_summary_2045_FINAL.xlsx", index=False)

with pd.ExcelWriter("scenario_all_outputs_FINAL.xlsx", engine="openpyxl") as
writer:
    scenario_params_df.to_excel(writer, sheet_name="Parameters", index=False)
    scenario_summary_all.to_excel(writer, sheet_name="Summary_all", index=False)
    scenario_population_export.to_excel(writer, sheet_name="Population",
index=False)
    scenario_fertility_export.to_excel(writer, sheet_name="Fertility",
index=False)

```

```

    scenario_mortality_deaths_export.to_excel(writer, sheet_name="Deaths",
index=False)
    scenario_migration_export.to_excel(writer, sheet_name="Migration",
index=False)
    scenario_asmr_age_export.to_excel(writer, sheet_name="ASMR_age_sex",
index=False)
    e0_scenario_new_wide.to_excel(writer, sheet_name="Life_expectancy",
index=False)

print("Сохранены файлы сценарного анализа:")
print("scenario_parameters_FINAL.xlsx")
print("scenario_population_total_FINAL.xlsx")
print("scenario_fertility_TFR_births_FINAL.xlsx")
print("scenario_mortality_deaths_FINAL.xlsx")
print("scenario_migration_net_FINAL.xlsx")
print("scenario_mortality_ASMR_by_age_sex_FINAL.xlsx")
print("scenario_life_expectancy_new_mortality_trend_long.xlsx")
print("scenario_life_expectancy_new_mortality_trend_wide.xlsx")
print("scenario_summary_all_indicators_FINAL.xlsx")
print("scenario_summary_2045_FINAL.xlsx")
print("scenario_all_outputs_FINAL.xlsx")

print("Сводка по сценариям на 2045 год:")
display(summary_2045_all)

```

```

# =====
# Проверка чувствительности к выбору исторического окна калибровки
# =====

import numpy as np
import pandas as pd
import warnings

warnings.filterwarnings("ignore")

# -----
# 1. Параметры проверки
# -----

WINDOW_STARTS = [1990, 1995, 2000, 2005]
HORIZON = 5
DRAWS = 1000
ORIGINS = tuple(range(2014, 2019))
MIN_TRAIN_YEARS = 10

np.random.seed(123)

# -----
# 2. Проверка наличия необходимых объектов
# -----

required_objects = [
    "pop", "mort", "fert", "mig",
    "AGE_LABELS", "REPRO_AGE_GROUPS",
    "AGE_TRANSFER",
    "SRB",
    "MIG_GROUP_CAP", "MIG_TRANSFORM", "MIG_SCALE", "RESID_INFLATE_TAU",

```

```

    "pca_arima_any", "surv_prob_exp_by_age", "births_from_asfr",
    "shape_migration_sex"
]

missing = [x for x in required_objects if x not in globals()]

if missing:
    raise NameError(
        "Не найдены необходимые объекты. Сначала необходимо выполнить основные ячейки модели. "
        f"Отсутствуют: {missing}"
    )

AGE_LABELS = [str(x) for x in AGE_LABELS]
REPRO_AGE_GROUPS = [str(x) for x in REPRO_AGE_GROUPS]

A_BACKTEST = np.asarray(AGE_TRANSFER, dtype=float).copy()

assert A_BACKTEST.shape == (len(AGE_LABELS), len(AGE_LABELS)), (
    f"A_BACKTEST имеет размер {A_BACKTEST.shape}, "
    f"ожидалось {(len(AGE_LABELS), len(AGE_LABELS))}"
)

p_male = SRB / (1 + SRB)
p_fem = 1 - p_male

# -----
# 3. Вспомогательные функции
# -----

def wide_window(df, value, sex=None, start_year=None, end_year=None,
columns_order=None):
    """
    Формирует wide-таблицу по заданному историческому окну.
    Для смертности и миграции sex принимает значения 'F' или 'M'.
    Для рождаемости sex можно оставить None.
    """
    d = df.copy()

    if start_year is not None:
        d = d.loc[d["Year"] >= int(start_year)].copy()

    if end_year is not None:
        d = d.loc[d["Year"] <= int(end_year)].copy()

    if sex is not None:
        d = d.loc[d["Sex"] == sex].copy()

    w = d.pivot_table(
        index="Year",
        columns="Age_Group",
        values=value,
        aggfunc="mean",
        observed=False
    ).sort_index()

    w.columns = [str(c) for c in w.columns]

    if columns_order is not None:

```

```

        w = w.reindex(columns=columns_order)

w = w.astype(float).interpolate(axis=0, limit_direction="both")
w = w.ffill().bfill()

return w

def run_backtest_for_window(window_start, origins=ORIGINS, horizon=HORIZON,
draws=DRAWS):
    """
    Для заданного исторического окна модели обучаются на данных
    window_start...origin,
    после чего строится прогноз на horizon лет и проверяется попадание
    фактической
    численности населения в 90% и 75% предиктивные интервалы.
    """
    results = []

    max_fact_year = int(pop["Year"].max())

    for origin in origins:
        origin = int(origin)

        if origin + horizon > max_fact_year:
            continue

        if origin - window_start + 1 < MIN_TRAIN_YEARS:
            continue

        mortF_w = wide_window(
            mort,
            "Mortality_Rate",
            sex="F",
            start_year=window_start,
            end_year=origin,
            columns_order=AGE_LABELS
        )

        mortM_w = wide_window(
            mort,
            "Mortality_Rate",
            sex="M",
            start_year=window_start,
            end_year=origin,
            columns_order=AGE_LABELS
        )

        fertF_w = wide_window(
            fert.query("Sex == 'F'"),
            "Fertility_Rate",
            sex=None,
            start_year=window_start,
            end_year=origin
        )

        migF_in = wide_window(
            mig,
            "Migration_Inflow",

```

```

        sex="F",
        start_year=window_start,
        end_year=origin,
        columns_order=AGE_LABELS
    )

migF_out = wide_window(
    mig,
    "Migration_Outflow",
    sex="F",
    start_year=window_start,
    end_year=origin,
    columns_order=AGE_LABELS
)

migM_in = wide_window(
    mig,
    "Migration_Inflow",
    sex="M",
    start_year=window_start,
    end_year=origin,
    columns_order=AGE_LABELS
)

migM_out = wide_window(
    mig,
    "Migration_Outflow",
    sex="M",
    start_year=window_start,
    end_year=origin,
    columns_order=AGE_LABELS
)

migF_w = migF_in - migF_out
migM_w = migM_in - migM_out

base_pop = pop.query("Year == @origin").copy()

base_F = (
    base_pop.query("Sex == 'F'")
    .pivot_table(index="Age_Group", values="Population", aggfunc="sum")
    .reindex(AGE_LABELS)
    .fillna(0.0)
    .values
    .ravel()
    .astype(float)
)

base_M = (
    base_pop.query("Sex == 'M'")
    .pivot_table(index="Age_Group", values="Population", aggfunc="sum")
    .reindex(AGE_LABELS)
    .fillna(0.0)
    .values
    .ravel()
    .astype(float)
)

mortF_m = pca_arima_any(

```

```

        mortF_w,
        steps=horizon,
        draws=draws,
        transform="log",
        tau=RESID_INFLATE_TAU
    )

mortM_m = pca_arima_any(
    mortM_w,
    steps=horizon,
    draws=draws,
    transform="log",
    tau=RESID_INFLATE_TAU
)

fert_m = pca_arima_any(
    fertF_w,
    steps=horizon,
    draws=draws,
    transform="log",
    tau=RESID_INFLATE_TAU
)

migF_m = pca_arima_any(
    migF_w,
    steps=horizon,
    draws=draws,
    transform=MIG_TRANSFORM,
    scale=MIG_SCALE,
    tau=RESID_INFLATE_TAU
)

migM_m = pca_arima_any(
    migM_w,
    steps=horizon,
    draws=draws,
    transform=MIG_TRANSFORM,
    scale=MIG_SCALE,
    tau=RESID_INFLATE_TAU
)

total_sims = np.zeros((draws, horizon))

for d in range(draws):
    popF = base_F.copy()
    popM = base_M.copy()

    for t in range(horizon):
        mortF_t = mortF_m["sims"][d, t, :]
        mortM_t = mortM_m["sims"][d, t, :]
        migF_t = migF_m["sims"][d, t, :]
        migM_t = migM_m["sims"][d, t, :]
        asfr_t = fert_m["sims"][d, t, :]

        survF = surv_prob_exp_by_age(mortF_t)
        survM = surv_prob_exp_by_age(mortM_t)

        popF = A_BACKTEST @ (popF * survF)
        popM = A_BACKTEST @ (popM * survM)

```

```

        popF += shape_migration_sex(migF_t, "F", cap=MIG_GROUP_CAP)
        popM += shape_migration_sex(migM_t, "M", cap=MIG_GROUP_CAP)

        B = births_from_asfr(popF, asfr_t)

        popF[0] += B * p_fem
        popM[0] += B * p_male

        total_sims[d, t] = popF.sum() + popM.sum()

years_check = list(range(origin + 1, origin + 1 + horizon))

fact = (
    pop.query("Year in @years_check")
    .groupby("Year")["Population"]
    .sum()
    .reindex(years_check)
    .values
)

for k, year in enumerate(years_check):
    q5, q95 = np.percentile(total_sims[:, k], [5, 95])
    q12, q88 = np.percentile(total_sims[:, k], [12.5, 87.5])
    med = np.median(total_sims[:, k])

    results.append({
        "Window_start": window_start,
        "Window": f"{window_start}-2023",
        "Origin": origin,
        "Year": year,
        "Horizon": year - origin,
        "Fact_Total": fact[k],
        "Median": med,
        "P5": q5,
        "P95": q95,
        "P12_5": q12,
        "P87_5": q88,
        "Covered90": (q5 <= fact[k] <= q95),
        "Covered75": (q12 <= fact[k] <= q88),
        "Width90": q95 - q5,
        "Width75": q88 - q12,
    })

return pd.DataFrame(results)

# -----
# 4. Запуск проверки по историческим окнам
# -----

all_backtests = []

for ws in WINDOW_STARTS:
    print(f"Расчёт окна {ws}-2023")
    bt = run_backtest_for_window(ws)
    all_backtests.append(bt)

bt_all = pd.concat(all_backtests, ignore_index=True)

```

```

# -----
# 5. Сводная таблица чувствительности
# -----

summary_rows = []

for ws, g in bt_all.groupby("Window_start"):
    cov90 = g["Covered90"].mean()
    cov75 = g["Covered75"].mean()

    by_h = g.groupby("Horizon").agg(
        Coverage90=("Covered90", "mean"),
        Coverage75=("Covered75", "mean"),
        Width90_mean=("Width90", "mean"),
        Width75_mean=("Width75", "mean"),
    ).reset_index()

    k90 = np.sqrt(0.90 / np.clip(by_h["Coverage90"], 1e-6, None))
    k75 = np.sqrt(0.75 / np.clip(by_h["Coverage75"], 1e-6, None))

    k90 = np.clip(k90, 1.0, 3.5)
    k75 = np.clip(k75, 1.0, 3.5)

    summary_rows.append({
        "Окно калибровки": f"{ws}-2023",
        "Число стартов backtesting": g["Origin"].nunique(),
        "Число проверок": len(g),
        "Среднее покрытие 90% интервала": round(cov90, 3),
        "Среднее покрытие 75% интервала": round(cov75, 3),
        "Средняя ширина 90% интервала, тыс. чел.": round(g["Width90"].mean() /
1000, 1),
        "Средняя ширина 75% интервала, тыс. чел.": round(g["Width75"].mean() /
1000, 1),
        "Средний k для 90% интервала": round(float(np.mean(k90)), 2),
        "Средний k для 75% интервала": round(float(np.mean(k75)), 2),
    })

summary_table = pd.DataFrame(summary_rows)

by_horizon = (
    bt_all
    .groupby(["Window", "Horizon"])
    .agg(
        Coverage90=("Covered90", "mean"),
        Coverage75=("Covered75", "mean"),
        Width90_mean=("Width90", "mean"),
        Width75_mean=("Width75", "mean"),
    )
    .reset_index()
)

by_horizon["Width90_mean_thousand"] = (by_horizon["Width90_mean"] /
1000).round(1)
by_horizon["Width75_mean_thousand"] = (by_horizon["Width75_mean"] /
1000).round(1)
by_horizon["Coverage90"] = by_horizon["Coverage90"].round(3)
by_horizon["Coverage75"] = by_horizon["Coverage75"].round(3)

```

```

print("Сводная таблица чувствительности к выбору исторического окна:")
display(summary_table)

print("Детализация по горизонту прогноза:")
display(by_horizon[
    [
        "Window",
        "Horizon",
        "Coverage90",
        "Coverage75",
        "Width90_mean_thousand",
        "Width75_mean_thousand"
    ]
])

# -----
# 6. Сохранение результатов
# -----

out_file = "appendix_window_sensitivity_backtesting.xlsx"

with pd.ExcelWriter(out_file, engine="openpyxl") as writer:
    summary_table.to_excel(writer, sheet_name="summary_for_appendix",
index=False)
    by_horizon.to_excel(writer, sheet_name="by_horizon", index=False)
    bt_all.to_excel(writer, sheet_name="all_backtest_results", index=False)

print(f"Файл сохранён: {out_file}")

```

ПРИЛОЖЕНИЕ 5

Таблица 1 – Входной массив возрастных коэффициентов смертности Санкт-Петербурга для возрастной группы 0-4 года, 1990-1998 гг., ‰

Год	Возрастная группа	Пол	Возрастной коэффициент смертности, ‰
1990	0-4	F	3,3
1991	0-4	F	3,15
1992	0-4	F	3,01
1993	0-4	F	2,88
1994	0-4	F	2,75
1995	0-4	F	2,63
1996	0-4	F	2,51
1997	0-4	F	2,4
1998	0-4	F	2,29
1990	0-4	M	4,38
1991	0-4	M	4,17
1992	0-4	M	3,97
1993	0-4	M	3,79
1994	0-4	M	3,61
1995	0-4	M	3,44
1996	0-4	M	3,28
1997	0-4	M	3,12
1998	0-4	M	2,98

Источник: составлено автором на основе входного массива возрастных коэффициентов смертности, использованного при формировании базы когортно-компонентного прогноза.

Примечание: F – женщины, M – мужчины. Показатели приведены в расчёте на 1000 человек соответствующего пола и возраста. Значения за 1990-1998 гг. для возрастной группы 0-4 года включены во входной массив как восстановленный фрагмент ряда.

Таблица 2 – Восстановленный массив возрастно-половых коэффициентов смертности Санкт-Петербурга, 1990-2010 гг.,
‰

А) Женщины

Возрастная группа	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
0-4	3,3	3,2	3	2,9	2,8	2,6	2,5	2,4	2,3	2,2	2,4	2,1	1,9	2	1,8	1,6	1,3	1,2	1,1	1,4	1,5
5-9	0,3	0,3	0,3	0,4	0,3	0,4	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,2	0,2	0,2	0,2	0,2
10-14	0,3	0,3	0,3	0,4	0,3	0,4	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,2	0,3	0,3
15-19	0,6	0,7	0,7	0,8	0,9	0,9	0,8	0,7	0,8	0,8	0,8	0,7	0,7	0,7	0,7	0,7	0,6	0,6	0,6	0,5	0,5
20-24	0,6	0,6	0,7	0,8	0,9	0,9	0,8	0,8	0,8	1	1	1	0,9	0,9	0,8	0,9	0,8	0,8	0,7	0,6	0,6
25-29	0,6	0,7	0,8	1	1,1	1,1	1	0,9	1	1	1,1	1,1	1,2	1,2	1,3	1,3	1,3	1,2	1,2	1,1	1,1
30-34	0,8	0,9	1	1,3	1,5	1,5	1,3	1,2	1,2	1,4	1,5	1,5	1,6	1,7	1,8	1,8	1,7	1,6	1,6	1,5	1,6
35-39	1,3	1,3	1,6	2	2,3	2,2	1,9	1,7	1,7	1,9	2	2,1	2,3	2,4	2,4	2,5	2,2	2,1	2,1	2	2
40-44	2	2,1	2,4	3,1	3,6	3,4	2,9	2,5	2,4	2,8	2,9	3	3,3	3,4	3,4	3,4	3	2,8	2,8	2,6	2,6
45-49	3,3	3,3	3,7	4,8	5,6	5,3	4,6	4	3,9	4,3	4,5	4,8	5	5,1	4,9	4,9	4,3	4	3,9	3,8	3,6
50-54	4,8	4,9	5,5	7,1	8,2	7,7	6,7	5,9	5,6	6,4	6,8	7,2	7,4	7,7	7,2	7,1	6,2	5,8	5,6	5,3	5,2
55-59	7,7	7,7	8	9,7	10,9	10,2	9,4	8,8	8,7	9,3	9,9	10,1	10,4	10,4	10,2	9,9	9	8,3	8,4	7,9	7,7
60-64	12,5	12,6	13,2	15,4	16,8	15,7	14,6	13,7	12,9	13,7	14,1	14,6	15,1	15,1	14,5	14	12,8	11,7	11,4	11,2	11,3
65-69	20,3	20,1	20,6	23,3	24,6	23,6	22,4	22	21,6	22,5	22,8	22,6	22,6	21,8	20,7	20,1	18,8	18,2	18	17,3	17,1
70-74	33,5	33,1	33,8	37,1	38,1	36,6	35,5	35	35,1	36,3	36,4	36,4	37	36,7	34,9	33,5	31,6	29,4	27,8	27,2	27,1
75-79	55,5	55,1	56,1	61,1	63,4	59,7	57,7	57,8	56,8	58,5	58,2	58,3	59,6	59,4	56,6	55,4	53,5	51,4	50,7	48,5	48,5
80-84	93,5	93,3	92,5	101,3	105,1	102,3	99,1	98,9	97,3	101,9	99,8	98,4	99,7	97,5	92	89,9	87,9	85,8	85,9	82,5	84,3
85+	168,4	166,2	169,6	189	189,9	187,8	180,1	180,4	177,9	187,9	188,3	186,6	190,9	188,4	176,2	174	168	162	160,2	154,3	156,3

Б) Мужчины

Возрастная группа	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
0-4	4,4	4,2	4	3,8	3,6	3,4	3,3	3,1	3	3	2,8	3	2,5	2,6	2,2	1,9	1,5	1,3	1,5	1,7	1,6
5-9	0,5	0,6	0,5	0,5	0,5	0,5	0,4	0,4	0,4	0,5	0,4	0,4	0,4	0,4	0,4	0,3	0,3	0,3	0,2	0,3	0,3
10-14	0,5	0,5	0,5	0,5	0,5	0,5	0,4	0,4	0,4	0,4	0,4	0,4	0,4	0,4	0,4	0,4	0,4	0,3	0,3	0,3	0,3
15-19	1,4	1,5	1,6	1,9	1,9	2,2	1,9	1,7	1,7	1,7	1,8	1,6	1,5	1,4	1,4	1,4	1,4	1,2	1,1	1	0,9
20-24	1,6	1,7	2,1	2,6	2,8	3	2,9	2,7	2,8	3,2	3,6	3	2,7	2,6	2,7	2,7	2,7	2,2	2	1,7	1,6
25-29	2,1	2,2	2,7	3,4	3,6	3,6	3,3	2,9	3	3,4	4	3,9	3,7	3,9	4	4,3	4,3	3,7	3,4	3	2,9
30-34	2,9	3,1	3,9	5,1	5,7	5,4	4,8	4,1	4,2	4,7	5,1	5,1	5,2	5,4	5,6	5,9	5,9	5,3	5,3	4,9	4,9
35-39	4,1	4,4	5,4	7,4	8,5	8	6,8	5,9	5,7	6,5	7,1	7,3	7,6	7,9	7,9	7,8	7,8	6,4	6,2	5,8	6
40-44	5,8	6,2	7,7	10,9	12,5	11,7	9,9	8,4	8,2	9,3	10,3	10,7	11,3	11,7	11,4	11,3	11,3	8,8	8,5	7,5	7,5
45-49	8,8	8,8	10,5	14,2	16,7	15,6	13,5	11,4	11,3	12,8	14,2	15	15,6	15,9	15,4	15,1	15,1	11,8	11,4	10,4	10,2
50-54	12,5	12,8	15,3	20,3	23,6	22,3	18,9	15,9	15,3	17,7	19,5	20,7	21,6	22,2	21,2	21	21	16,4	16	14,9	14,4
55-59	18,5	18,5	20,1	25,4	29,6	27,8	24,8	22,8	22,2	24,7	26,5	27,1	27,5	27,5	26,8	26,3	26,3	21,7	22	20,6	20,4
60-64	28,1	28,4	30,4	37,5	42,2	39	35	32,1	30,6	34,2	36,1	37,7	39,8	39,8	37,4	36	36	29,9	29,7	29,3	29,3
65-69	38	38,7	40,3	48,5	52,3	50,3	47	45,6	44,6	47,4	48,5	49	49,5	48,7	46,9	45,9	45,9	41,6	41,7	39,8	39,4
70-74	55	54,6	56,2	64	66,2	62,8	60,1	59,5	59,2	63,5	65,5	66,1	67,7	67,4	64,4	62,9	62,9	55,1	54	53,8	53,8
75-79	80,3	80,5	80,7	90,9	93,3	86,9	82,2	81,3	79,1	81,2	83,1	83	85,7	87,4	85,7	85,3	85,3	79,6	78,5	75,8	75,1
80-84	119,1	118,9	117,8	132	134,1	126,9	122,3	121	117,6	123,5	121,6	118,8	120,1	117,8	112,8	112	112	107,2	109	106,3	108,9
85+	176,3	177,5	181,6	201,7	201,8	188	178,7	177,3	172,2	183,4	184	181,5	184,8	183,9	174,1	172,1	172,1	162,3	157,6	150,9	150,4

Источник: составлено автором на основе восстановленного массива возрастно-половых коэффициентов смертности Санкт-Петербурга, использованного при формировании базы когортно-компонентного прогноза.

Примечание: показатели приведены в расчёте на 1000 человек соответствующего пола и возраста и округлены до одного десятичного знака.

Таблица 1 – Восстановленные миграционные потоки Санкт-Петербурга по полу и возрасту, 1990-2002 гг., тыс. чел.

а) Женщины, прибывшие

Возраст	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
0-4	0,45	0,473	0,496	0,521	0,547	0,575	0,603	0,633	0,665	0,699	0,733	0,77	0,809
5-9	1,092	1,072	1,053	1,034	1,015	0,995	0,976	0,957	0,937	0,918	0,899	0,88	0,86
10-14	1,815	1,76	1,706	1,651	1,597	1,542	1,488	1,434	1,379	1,325	1,27	1,216	1,161
15-19	1,728	1,741	1,757	1,777	1,802	1,834	1,873	1,922	1,984	2,062	2,159	2,28	2,432
20-24	2,196	2,023	1,895	1,813	1,775	1,783	1,835	1,933	2,076	2,264	2,497	2,776	3,099
25-29	8,865	7,881	6,986	6,179	5,461	4,83	4,288	3,834	3,468	3,191	3,002	2,901	2,889
30-34	7,417	6,644	5,931	5,277	4,682	4,146	3,67	3,252	2,894	2,596	2,357	2,176	2,056
35-39	1,374	1,375	1,376	1,377	1,379	1,382	1,385	1,39	1,397	1,406	1,419	1,436	1,46
40-44	1,455	1,455	1,455	1,456	1,456	1,457	1,458	1,459	1,461	1,463	1,467	1,472	1,479
45-49	1,083	1,083	1,083	1,083	1,084	1,084	1,085	1,086	1,087	1,089	1,092	1,095	1,101
50-54	0,523	0,549	0,576	0,605	0,635	0,666	0,699	0,734	0,77	0,808	0,849	0,891	0,935
55-59	0,314	0,338	0,363	0,389	0,418	0,449	0,482	0,517	0,556	0,597	0,641	0,688	0,739
60-64	0,433	0,451	0,47	0,49	0,511	0,532	0,555	0,579	0,603	0,629	0,655	0,683	0,712
65-69	0,963	0,972	0,981	0,991	1	1,009	1,019	1,028	1,038	1,047	1,057	1,067	1,077
70-74	0,629	0,635	0,641	0,647	0,653	0,659	0,665	0,671	0,677	0,683	0,69	0,696	0,703
75-79	0,481	0,485	0,49	0,494	0,499	0,504	0,508	0,513	0,518	0,522	0,527	0,532	0,537
80-84	0,416	0,42	0,423	0,426	0,43	0,433	0,436	0,44	0,443	0,447	0,45	0,454	0,457
85+	0,254	0,257	0,26	0,263	0,266	0,269	0,272	0,275	0,278	0,281	0,284	0,287	0,291

б) Женщины, выбывшие

Возраст	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
0-4	0,394	0,418	0,443	0,467	0,492	0,516	0,541	0,566	0,59	0,615	0,639	0,664	0,688
5-9	0,732	0,722	0,711	0,701	0,69	0,68	0,669	0,658	0,648	0,637	0,627	0,616	0,606
10-14	1,091	1,058	1,025	0,993	0,96	0,927	0,894	0,861	0,828	0,795	0,763	0,73	0,697
15-19	1,694	1,668	1,643	1,617	1,592	1,566	1,541	1,515	1,49	1,465	1,439	1,414	1,388
20-24	8,941	8,552	8,163	7,774	7,385	6,996	6,607	6,218	5,829	5,44	5,05	4,661	4,272
25-29	1,545	1,569	1,593	1,617	1,641	1,665	1,689	1,713	1,737	1,761	1,785	1,809	1,833
30-34	1,358	1,357	1,356	1,354	1,353	1,352	1,351	1,349	1,348	1,347	1,346	1,344	1,343
35-39	0,926	0,934	0,941	0,949	0,957	0,964	0,972	0,979	0,987	0,995	1,002	1,01	1,017
40-44	1,614	1,578	1,542	1,507	1,471	1,435	1,399	1,363	1,327	1,292	1,256	1,22	1,184
45-49	1,256	1,225	1,195	1,164	1,133	1,102	1,072	1,041	1,01	0,979	0,949	0,918	0,887
50-54	1,052	1,038	1,025	1,012	0,999	0,985	0,972	0,959	0,946	0,932	0,919	0,906	0,893
55-59	0,75	0,753	0,756	0,76	0,763	0,766	0,77	0,773	0,777	0,78	0,784	0,787	0,79
60-64	0,764	0,758	0,753	0,748	0,742	0,737	0,732	0,727	0,721	0,716	0,711	0,706	0,701
65-69	1,282	1,255	1,229	1,203	1,177	1,152	1,128	1,104	1,081	1,058	1,035	1,013	0,992
70-74	0,781	0,764	0,748	0,733	0,717	0,702	0,687	0,673	0,659	0,645	0,631	0,618	0,605
75-79	0,605	0,592	0,579	0,567	0,555	0,543	0,531	0,52	0,509	0,498	0,488	0,477	0,467
80-84	0,52	0,508	0,497	0,485	0,474	0,463	0,452	0,442	0,432	0,422	0,412	0,402	0,393
85+	0,311	0,306	0,3	0,294	0,289	0,283	0,278	0,273	0,268	0,263	0,258	0,253	0,249

в) Мужчины, прибывшие

Возраст	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
0-4	1,178	1,158	1,138	1,118	1,098	1,078	1,058	1,038	1,018	0,998	0,978	0,958	0,938
5-9	1,163	1,14	1,118	1,095	1,073	1,05	1,028	1,006	0,983	0,961	0,938	0,916	0,893
10-14	1,876	1,82	1,764	1,708	1,652	1,596	1,54	1,484	1,428	1,372	1,316	1,26	1,204
15-19	2,503	2,503	2,503	2,503	2,503	2,504	2,504	2,505	2,507	2,509	2,514	2,522	2,536
20-24	9,955	8,86	7,86	6,952	6,137	5,416	4,789	4,254	3,813	3,465	3,21	3,048	2,98
25-29	9,685	8,649	7,696	6,828	6,045	5,346	4,731	4,2	3,754	3,392	3,114	2,921	2,812
30-34	9,333	8,393	7,519	6,712	5,971	5,296	4,688	4,147	3,672	3,263	2,921	2,645	2,436
35-39	1,611	1,611	1,611	1,612	1,612	1,613	1,614	1,616	1,619	1,623	1,629	1,638	1,651
40-44	1,455	1,455	1,455	1,455	1,455	1,456	1,456	1,457	1,458	1,46	1,463	1,466	1,472
45-49	1,062	1,063	1,063	1,063	1,063	1,064	1,065	1,066	1,067	1,07	1,073	1,078	1,084
50-54	0,391	0,413	0,436	0,461	0,487	0,515	0,544	0,575	0,607	0,642	0,678	0,716	0,757
55-59	0,203	0,22	0,238	0,257	0,278	0,301	0,326	0,353	0,382	0,414	0,448	0,485	0,525
60-64	0,166	0,178	0,19	0,204	0,218	0,234	0,25	0,268	0,287	0,308	0,33	0,353	0,378
65-69	0,16	0,17	0,181	0,192	0,204	0,216	0,23	0,244	0,259	0,275	0,292	0,31	0,33
70-74	0,141	0,147	0,152	0,158	0,164	0,17	0,177	0,183	0,19	0,197	0,205	0,213	0,221
75-79	0,153	0,155	0,156	0,158	0,16	0,161	0,163	0,165	0,166	0,168	0,17	0,171	0,173
80-84	0,191	0,188	0,184	0,181	0,178	0,175	0,171	0,168	0,165	0,162	0,159	0,157	0,154
85+	0,265	0,245	0,226	0,208	0,192	0,177	0,163	0,15	0,138	0,127	0,117	0,108	0,1

г) Мужчины, выбывшие

Возраст	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
0-4	0,585	0,602	0,619	0,637	0,654	0,671	0,688	0,705	0,722	0,739	0,757	0,774	0,791
5-9	0,951	0,932	0,912	0,892	0,872	0,853	0,833	0,813	0,794	0,774	0,754	0,734	0,715
10-14	1,213	1,176	1,139	1,102	1,065	1,028	0,991	0,954	0,917	0,88	0,843	0,806	0,769
15-19	3,692	3,529	3,366	3,203	3,041	2,878	2,715	2,552	2,389	2,226	2,063	1,9	1,737
20-24	9,463	9,045	8,627	8,209	7,79	7,372	6,954	6,536	6,118	5,7	5,282	4,863	4,445
25-29	3,298	3,214	3,129	3,045	2,96	2,876	2,791	2,707	2,622	2,538	2,454	2,369	2,285
30-34	2,315	2,264	2,214	2,163	2,113	2,062	2,012	1,961	1,911	1,86	1,81	1,759	1,709
35-39	1,513	1,493	1,473	1,453	1,433	1,414	1,394	1,374	1,354	1,334	1,314	1,295	1,275
40-44	1,865	1,825	1,784	1,744	1,703	1,663	1,623	1,582	1,542	1,501	1,461	1,42	1,38
45-49	1,294	1,27	1,246	1,223	1,199	1,175	1,151	1,127	1,104	1,08	1,056	1,032	1,008
50-54	0,96	0,961	0,962	0,962	0,963	0,964	0,965	0,966	0,967	0,968	0,968	0,969	0,97
55-59	0,597	0,602	0,607	0,612	0,616	0,621	0,626	0,631	0,636	0,641	0,646	0,651	0,656
60-64	0,472	0,474	0,476	0,478	0,48	0,482	0,484	0,487	0,489	0,491	0,493	0,496	0,498
65-69	0,292	0,298	0,304	0,31	0,317	0,323	0,33	0,337	0,344	0,351	0,358	0,366	0,373
70-74	0,23	0,23	0,23	0,231	0,231	0,231	0,231	0,231	0,231	0,231	0,232	0,232	0,232
75-79	0,286	0,276	0,266	0,257	0,248	0,239	0,231	0,222	0,214	0,207	0,199	0,192	0,186
80-84	0,386	0,359	0,335	0,312	0,29	0,27	0,252	0,234	0,218	0,203	0,189	0,176	0,164
85+	0,542	0,473	0,413	0,361	0,315	0,275	0,24	0,21	0,183	0,16	0,14	0,122	0,107

Источник: составлено автором на основе восстановленного входного массива миграционных потоков, использованного при формировании базы когортно-компонентного прогноза.

Примечание: показатели приведены в тыс. человек и округлены до 0,001 тыс. чел. В таблице представлен восстановленный участок миграционного ряда за 1990-2002 гг.

Таблица 2 – Миграционные потоки Санкт-Петербурга в возрастных группах 80-84 и 85+, 2003-2023 гг., тыс. чел.

Год	Пол	Прибывшие 80-84	Прибывшие 85+	Выбывшие 80-84	Выбывшие 85+
2003	женщины	0,461	0,294	0,384	0,244
2003	мужчины	0,151	0,092	0,153	0,093
2004	женщины	0,463	0,269	0,351	0,205
2004	мужчины	0,156	0,089	0,139	0,079
2005	женщины	0,437	0,241	0,318	0,166
2005	мужчины	0,153	0,078	0,137	0,07
2006	женщины	0,448	0,238	0,309	0,165
2006	мужчины	0,159	0,072	0,141	0,064
2007	женщины	0,405	0,225	0,284	0,158
2007	мужчины	0,149	0,06	0,121	0,048
2008	женщины	0,396	0,236	0,255	0,152
2008	мужчины	0,101	0,038	0,072	0,027
2009	женщины	0,338	0,212	0,229	0,144
2009	мужчины	0,083	0,031	0,063	0,024
2010	женщины	0,487	0,318	0,326	0,214
2010	мужчины	0,133	0,052	0,093	0,036
2011	женщины	0,689	0,473	0,625	0,429
2011	мужчины	0,182	0,076	0,199	0,083
2012	женщины	0,825	0,598	0,666	0,483
2012	мужчины	0,251	0,12	0,214	0,102
2013	женщины	0,973	0,759	0,749	0,585
2013	мужчины	0,268	0,151	0,223	0,125
2014	женщины	1,054	0,899	0,811	0,692
2014	мужчины	0,294	0,188	0,203	0,129
2015	женщины	1,069	0,975	0,789	0,72
2015	мужчины	0,315	0,218	0,247	0,171
2016	женщины	1,103	1,028	0,859	0,8
2016	мужчины	0,379	0,254	0,27	0,18
2017	женщины	1,317	1,139	0,926	0,801
2017	мужчины	0,441	0,257	0,32	0,186
2018	женщины	1,31	0,995	1,051	0,799
2018	мужчины	0,475	0,248	0,345	0,18
2019	женщины	1,351	0,916	1,006	0,681
2019	мужчины	0,481	0,241	0,378	0,19
2020	женщины	1,397	0,917	0,996	0,653
2020	мужчины	0,443	0,227	0,346	0,178
2021	женщины	1,329	0,665	0,999	0,52
2021	мужчины	0,431	0,293	0,303	0,207
2022	женщины	1,174	0,914	0,918	0,715

2022	мужчины	0,369	0,231	0,317	0,199
2023	женщины	1,013	1,068	0,693	0,731
2023	мужчины	0,354	0,306	0,242	0,208

Источник: составлено автором на основе входного массива миграционных потоков, использованного при формировании базы когортно-компонентного прогноза.

Примечание: показатели приведены в тыс. человек и округлены до 0,001 тыс. чел.

Таблица отражает результат деления открытой возрастной группы 80+ на подгруппы 80-84 и 85+ для миграционных потоков.

Итоговый вероятностный прогноз численности населения Санкт-Петербурга

Таблица 1 – Итоговый вероятностный прогноз численности населения Санкт-Петербурга, 2024-2045 гг., тыс. чел.

Год	Медиана	Нижняя граница 90% интервала	Верхняя граница 90% интервала	Нижняя граница 75% интервала	Верхняя граница 75% интервала
2024	5636,5	5590,3	5716,6	5593,3	5698,8
2025	5675,3	5587,4	5791,7	5567,4	5796,6
2026	5707,1	5585,9	5854,7	5602,4	5824
2027	5733,5	5551,1	5944	5588	5892,1
2028	5754,5	5506	6029,4	5593,5	5927,3
2029	5770,5	5503,6	6066,4	5597,7	5958,9
2030	5782,1	5498,1	6095,6	5598	5980,9
2031	5789,3	5488,2	6120,1	5597,4	5994,8
2032	5791,9	5477,6	6137,9	5593,5	6005,7
2033	5792,4	5467,8	6147,7	5585,3	6013,1
2034	5790	5454,1	6154,8	5574,4	6017,9
2035	5784,8	5441,7	6153,5	5564,1	6016,6
2036	5776,5	5425,8	6154,2	5551	6013,8
2037	5764,7	5408,2	6148,8	5535,2	6009,4
2038	5751,4	5388,7	6141,8	5516,8	6002,9
2039	5737	5372,3	6132,5	5498	5991,3
2040	5720,4	5350,7	6117	5478,3	5978
2041	5703	5323,4	6110,6	5459	5962,4
2042	5684	5305,2	6091,7	5437,5	5945,9
2043	5663,5	5277,6	6078,2	5413	5928,5
2044	5642,4	5251,1	6060,2	5391,5	5907,1
2045	5619,7	5225,3	6040,9	5368,5	5886,8

Источник: составлено автором на основе итогового массива калиброванных предиктивных интервалов вероятностного прогноза численности населения Санкт-Петербурга.

Примечание: показатели приведены в тыс. человек и округлены до 0,1 тыс. чел. Границы 90% интервала соответствуют значениям P5 и P95, границы 75% интервала – значениям P12,5 и P87,5.

Прогнозная возрастно-половая структура населения Санкт-Петербурга

Таблица 1 – Прогнозная возрастная структура женского населения Санкт-Петербурга по реперным годам, медиана, тыс. чел.

Возрастная группа	2024	2030	2035	2040	2045
0-4	121,7	119	112,7	108,1	104,7
5-9	138,5	126,9	120,2	114,2	109,6
10-14	132,2	132,6	127,3	121,2	115,6
15-19	128,1	150,8	151,9	146,7	139,6
20-24	147,8	144	149,4	149,6	145,3
25-29	150,9	160,7	166	169,6	169,2
30-34	201,4	180,8	179,8	182,7	184,8
35-39	263,8	214,5	197,5	192,7	192,9
40-44	259,1	253,5	228,5	211,2	203,1
45-49	229,7	261,7	251,5	231,8	215,9
50-54	209,6	240,1	251,1	243,1	227,7
55-59	198,4	214,8	233,4	239,6	232,7
60-64	221,8	203,7	212,3	223,2	225,8
65-69	216	203,5	198,6	203,5	209,1
70-74	185,6	194,8	187,3	184,1	186
75-79	124,4	159,9	162,5	158,5	156,5
80-84	80,3	109,1	120,3	120,7	118,3
85+	85,1	111,6	133,7	146,1	149,5

Таблица 2 – Прогнозная возрастная структура мужского населения Санкт-Петербурга по реперным годам, медиана, тыс. чел.

Возрастная группа	2024	2030	2035	2040	2045
0-4	127,9	125,9	119,3	114,5	110,9
5-9	146,3	134,1	127,3	121,1	116,3
10-14	139,3	140,1	134,9	128,8	123,1
15-19	127,9	147,8	146	139,6	132,5
20-24	165,7	148,4	148,3	145	139,2
25-29	146,4	160,1	162	161,9	158,9
30-34	183,2	172,5	174,8	176,2	176,1
35-39	241,5	198,5	188,7	186,7	186,9
40-44	229	219,4	201,7	192,2	188,3
45-49	197,6	216,5	208,6	196,4	187,9
50-54	177,8	195,7	201,9	195,7	186,7
55-59	155,9	170,6	182,2	185,1	180,2
60-64	151,7	149,5	157,1	163,8	164,5
65-69	128,7	129,7	132,4	137,6	141,2
70-74	96,4	105,5	107,5	110,4	113,7
75-79	58,1	75,7	80,7	83	85,3
80-84	31,4	45,9	52,9	56	57,7
85+	27,3	42,7	55,8	65,3	71

Источник: составлено автором на основе массива прогнозной возрастно-половой структуры населения Санкт-Петербурга, полученного в результате вероятностного когортно-компонентного моделирования.

Примечание: показатели приведены в тыс. человек и округлены до 0,1 тыс. чел.; представлены медианные значения прогнозного распределения.

ПРИЛОЖЕНИЕ 9

Прогнозные значения компонент воспроизводства населения Санкт-Петербурга

Таблица 1 – Прогнозные возрастные коэффициенты рождаемости Санкт-Петербурга по реперным годам, медиана, ‰

Возрастная группа	2024	2030	2035	2040	2045
15-19	8,5	9,6	10,3	10,8	11,3
20-24	47,2	50,2	52,1	53,5	54,7
25-29	91,2	87,1	84,6	82,9	81,7
30-34	80,9	71,5	66,2	62,6	60,3
35-39	44,7	37,1	33,1	30,6	28,9
40-44	11,1	8,8	7,6	6,9	6,4
45-49	1,1	0,7	0,5	0,4	0,4
50-54	0,1	0	0	0	0

Таблица 2 – Прогнозные возрастно-половые коэффициенты смертности Санкт-Петербурга по реперным годам, медиана, ‰

Возрастная группа	Пол	2024	2030	2035	2040	2045
0-4	женщины	0,89	1,01	1,1	1,17	1,24
0-4	мужчины	0,95	1,07	1,15	1,21	1,27
20-24	женщины	0,38	0,41	0,43	0,45	0,47
20-24	мужчины	0,97	1,01	1,05	1,1	1,13
40-44	женщины	2,11	2,17	2,21	2,24	2,27
40-44	мужчины	5,88	5,95	6,01	6,1	6,2
60-64	женщины	8,57	9,07	9,42	9,69	9,93
60-64	мужчины	22,39	22,82	23,17	23,56	23,91
80-84	женщины	67,01	70,02	72,04	73,61	75,03
80-84	мужчины	89,9	91,5	92,73	93,84	94,86
85+	женщины	140,54	144,6	147,29	149,35	151,17
85+	мужчины	131,36	133,93	135,89	137,65	139,23

Таблица 3 – Прогнозная миграционная компонента Санкт-Петербурга по полу и возрасту по реперным годам, медиана, чел.

а) женщины

Возрастная группа	2024	2030	2035	2040	2045
0-4	0	0	0	0	0
5-9	219	232	238	242	245
10-14	485	491	496	499	502
15-19	3 964	3 577	3 293	3 001	2 759
20-24	0	0	0	0	0
25-29	1 933	2 330	2 492	2 613	2 665
30-34	1 155	1 387	1 489	1 571	1 618
35-39	730	794	811	811	803
40-44	4 973	1 400	737	415	274
45-49	4 544	1 117	544	284	180
50-54	84	74	56	35	21
55-59	10	15	13	9	6
60-64	29	22	16	10	6
65-69	385	157	88	49	29
70-74	1 549	372	174	87	51
75-79	1 100	280	134	68	40
80-84	950	237	114	58	34
85+	732	187	93	48	30

б) мужчины

Возрастная группа	2024	2030	2035	2040	2045
0-4	0	0	0	0	0
5-9	310	252	246	242	238
10-14	600	500	502	508	507
15-19	2 458	391	113	49	29
20-24	0	0	0	0	0
25-29	1	129	160	168	198
30-34	0	126	159	167	201
35-39	0	63	66	62	69
40-44	0	4	1	0	0
45-49	7	14	4	1	1
50-54	3	1	0	0	0
55-59	6	1	0	0	0
60-64	114	1	0	0	0
65-69	163	1	0	0	0
70-74	174	1	0	0	0
75-79	134	1	0	0	0
80-84	102	1	0	0	0
85+	59	1	0	0	0

Таблица 4 – Прогнозные значения родившихся, умерших и чистой миграции в Санкт-Петербурге, 2024-2045 гг., тыс. чел.

Год	Родившиеся, медиана	Родившиеся, P5-P95	Умершие, медиана	Умершие, P5-P95	Чистая миграция, медиана	Чистая миграция, P5-P95
2024	53,3	43,7-70,9	62,1	58,5-66,0	46	18,8-107,0
2025	51,5	42,6-68,1	65	61,1-69,4	43	18,5-103,9
2026	50	41,6-65,9	67,6	63,4-72,4	40,5	18,4-99,8
2027	48,7	40,7-63,7	70	65,7-74,9	40,2	18,2-98,1
2028	47,7	40,0-61,7	72,5	68,1-77,7	38	18,0-95,6
2029	46,8	39,5-60,1	74,9	70,3-80,3	36,4	17,9-94,2
2030	46,2	39,1-59,6	77,1	72,2-82,8	34	17,7-91,3
2031	45,5	38,6-58,1	79,2	74,3-84,8	32	17,5-85,7
2032	45	38,4-57,1	81,3	76,2-87,0	30,7	17,4-84,1
2033	44,6	38,0-56,2	83,1	77,9-88,9	30	17,3-79,9
2034	44,2	38,0-55,7	84,8	79,6-90,8	29	17,2-78,7
2035	43,9	37,8-54,9	86,4	81,1-92,4	28,5	17,1-77,7
2036	43,6	37,6-54,7	87,7	82,4-93,8	27,5	17,1-73,7
2037	43,5	37,5-54,0	89,1	83,5-95,2	26,9	17,0-73,2
2038	43,2	37,2-53,5	90,2	84,7-96,3	26,2	16,9-71,8
2039	43,1	37,2-53,2	91,2	85,7-97,4	25,4	16,9-69,8
2040	42,7	37,1-52,2	92	86,6-98,1	25,2	16,8-66,3
2041	42,7	36,9-52,3	92,9	87,2-99,1	24,7	16,8-66,3
2042	42,5	36,7-51,7	93,5	87,9-99,8	24,2	16,6-65,0
2043	42,1	36,5-51,3	94,2	88,5-100,3	24,1	16,6-63,1
2044	42	36,3-51,2	94,6	88,9-100,9	23,6	16,5-62,1
2045	41,6	36,1-50,9	95	89,4-101,2	23,4	16,6-61,2

Источник: составлено автором на основе массивов прогнозных компонент, полученных в результате вероятностного когортно-компонентного моделирования.

Примечание: в таблицах 1-4 представлены медианные значения прогнозного распределения; единицы измерения указаны в названиях таблиц.

Результаты сценарного прогноза численности населения Санкт-Петербурга

Таблица 1 – Медианные значения сценарного прогноза численности населения Санкт-Петербурга, 2024-2045 гг., тыс. чел.

Год	С1: благоприятный	С2: умеренно благоприятный	С3: умеренно неблагоприятный	С4: неблагоприятный
2024	5667,7	5656,1	5642,8	5625,4
2025	5739,2	5715,3	5688,2	5652,7
2026	5804,6	5768,3	5727,0	5673,0
2027	5865,7	5816,6	5760,6	5687,6
2028	5922,6	5860,1	5789,1	5696,5
2029	5975,5	5899,5	5813,0	5700,3
2030	6024,5	5934,6	5832,7	5699,2
2031	6070,3	5966,1	5848,2	5694,1
2032	6112,0	5993,6	5859,2	5684,0
2033	6152,7	6019,6	5868,5	5671,4
2034	6190,4	6042,6	5874,7	5656,2
2035	6226,6	6063,5	5878,5	5637,6
2036	6260,0	6081,5	5879,0	5615,6
2037	6290,3	6096,5	5876,4	5590,7
2038	6320,1	6110,5	5872,4	5563,5
2039	6347,9	6123,0	5867,2	5535,5
2040	6374,9	6134,1	5860,1	5505,8
2041	6401,0	6144,3	5852,1	5474,6
2042	6425,6	6152,8	5842,4	5441,5
2043	6449,7	6160,7	5831,4	5407,0
2044	6472,9	6167,4	5819,7	5372,5
2045	6494,6	6172,8	5806,4	5335,8

Таблица 2 – Сценарные значения ожидаемой продолжительности жизни при рождении в Санкт-Петербурге, медиана, лет

Год	С1 муж.	С1 жен.	С2 муж.	С2 жен.	С3 муж.	С3 жен.	С4 муж.	С4 жен.
2024	71,8	80,1	71,8	80,1	71,7	80	71,6	80
2025	71,9	80,1	71,8	80	71,7	79,9	71,5	79,8
2026	72,1	80,1	71,9	80	71,7	79,9	71,4	79,7
2030	72,6	80,2	72,2	79,9	71,7	79,6	71	79,3
2035	73,3	80,4	72,6	79,9	71,6	79,3	70,6	78,7
2040	73,9	80,6	72,9	80	71,6	79,1	70,1	78,3
2045	74,6	80,9	73,3	80,1	71,6	79	69,6	77,9

Источник: составлено автором на основе результатов сценарного вероятностного прогноза численности населения Санкт-Петербурга.