

УДК 314.8:519.237

DOI: <https://doi.org/10.17323/ej.2026.33626>

Прогнозирование численности населения современного мегаполиса (на примере Санкт-Петербурга): методологические аспекты

Рябчикова Д.А.

Статья посвящена вероятностному когортно-компонентному прогнозированию населения мегаполиса как открытой демографической системы, где миграция определяет динамику численности и возрастной структуры. В рассматриваемой методологии рождаемость, смертность и миграция моделируются как стохастические возрастно-половые процессы: возрастные профили редуцируются методом PCA, динамика временных коэффициентов описывается авторегрессионной моделью первого порядка AR(1) (эквивалент ARIMA(1,0,0)), а прогноз формируется как распределение траекторий в Монте-Карло-симуляции. Миграционная компонента задана через относительные возрастно-половые ставки с дрейфом и дополнительную уровневую корректировку (bump-компонента), отражающую возможный недоучет миграционного прироста; интервальные оценки валидируются ретроспективной проверкой (backtesting) и калибруются по эмпирическому покрытию. Все расчеты произведены по данным Санкт-Петербурга, эмпирическая база охватывает 1990–2023 гг., горизонт прогноза – до 2045 г., что обеспечивает сопоставимость с официальными оценками Росстата. Медианная траектория фиксирует переход города к режиму демографического «плато»: к 2045 г. численность составляет 5,62 млн человек при 90-процентном интервале 5,225 – 6,041 млн человек. Одновременно усиливается старение населения (доля жителей 65+ достигает 23% к 2045 г.), и роль миграции смещается от источника роста к механизму компенсации естественной убыли и стабилизации возрастной структуры. Сценарные прогнозы показывают, что долговременная развилка траекторий определяется прежде всего накопленным миграционным эффектом (к 2045 г. – от 5,32 млн человек до 6,40 млн в крайних вариантах). Сопоставление с прогнозом Росстата выявляет умеренные расхождения, траектория официального прогноза укладывается в доверительные интервалы вероятностной модели. Предложенный инструментарий может применяться в прогнозировании численности населения на уровне крупных городов и регионов страны.

Рябчикова Дарья Анатольевна – аспирант, Санкт-Петербургский государственный экономический университет. E-mail: d.sinitzyna@yandex.ru

Статья поступила: 26.01.2026/Статья принята: 10.03.2026.

Ключевые слова: стохастическая когортно-компонентная модель; Монте-Карло; РСА-AR(1); backtesting; миграция; открытая демографическая система; мегаполис; Санкт-Петербург; старение населения.

Для цитирования: Рябчикова Д.А. Прогнозирование численности населения современного мегаполиса (на примере Санкт-Петербурга): методологические аспекты. *Экономический журнал ВШЭ*. 2026; 30(1): 182–205.

For citation: Ryabchikova D.A. Population Forecasting in a Modern Metropolis (using St. Petersburg as a Case Study): Methodological Aspects. *HSE Economic Journal*. 2026; 30(1): 182–205. (In Russ.)

1. Введение

Демографическое развитие мегаполиса имеет особенности по сравнению с закономерностями страны. В крупных городах ключевым фактором формирования численности и возрастной структуры населения выступает миграция, а динамика рождаемости и смертности дополняет ее влияние. Мегаполис при этом представляет собой открытую демографическую систему. Если на национальном уровне внутренние миграционные притоки и оттоки в сумме компенсируют друг друга, то в мегаполисе такое равновесие не достигается; естественная убыль здесь редко возмещается рождаемостью, и именно устойчиво ненулевое миграционное сальдо становится основным механизмом изменения численности и структуры населения.

Перспективное развитие структуры населения имеет огромное значение, поскольку стратегические решения опираются на ожидаемый состав населения. Детерминистские демографические прогнозы, построенные как набор сценариев, задают несколько траекторий, соответствующих разным гипотезам о будущей рождаемости, смертности и миграции.

В официальной практике Росстата многовариантные сценарные прогнозы (низкий, средний и высокий варианты) публикуются для Российской Федерации в целом и сопровождаются оценками ключевых демографических показателей (численности населения, структуры по полу и возрасту, рождаемости, смертности, миграции, демографической нагрузки и др.). При представлении результатов по субъектам РФ (в том числе Санкт-Петербургу) в открытых публикациях приводится средний вариант.

Вместе с тем такие прогнозы обычно не предполагают вероятностной интерпретации полученных траекторий и, следовательно, не дают количественной оценки уровня неопределенности прогноза в виде предиктивных интервалов [Vanella, Deschermeier, 2020; Keilman et al., 2002]. Основная идея вероятностного прогноза заключается в том, чтобы «учесть главные источники неопределенности и оценить вероятность того, что фактическая численность населения не выйдет за установленные прогнозируемые пределы» [Никитина, Щербов, 2007]. Следует, однако, учитывать, что вероятностные прогнозы на основе временных рядов имеют методологические ограничения при описании реальных демографических процессов, в том числе из-за возможных структурных сдвигов, что требует взвешенной интерпретации результатов и проверок устойчивости¹ [Keilman et al., 2002].

¹ Orig. (Keilman): «There is no guarantee that the statistical model that we used to compute prediction intervals is the “real” or “correct” one. Unfortunately, different statistical models may produce widely

Прогнозирование численности населения мегаполиса осложняется высокой изменчивостью ключевых демографических компонент. Миграционные потоки крупных городов подвержены резким колебаниям, способным за сравнительно короткий период существенно изменить траекторию динамики, а возрастная структура формируется под воздействием выраженных когортных волн. Дополнительный источник неопределенности связан с расхождениями между различными системами учета: данными переписи населения и текущего учета. Например, численность населения Санкт-Петербурга по итогам ВПН-2020² составляла 5,602 млн человек в пересчете на 01.10.2021 г.³, тогда как по оценке текущего учета – 5,3715 млн человек на ту же дату, 01.10.2021 г.⁴ Так что даже исходный уровень численности населения выражается не единственным значением, а интервалом оценок в зависимости от метода учета. Это напрямую влияет как на выбор базовой точки прогноза, так и на ширину доверительных интервалов, поскольку неопределенность присутствует уже на уровне исходной численности.

Исходя из этих соображений оправдан переход от детерминированных демографических моделей к стохастическим, которые «интегрируют различные источники неопределенности в структуру моделирования» и создают «естественную вероятностную основу для прогнозирования будущих численностей населения» [Носова, 2019]. Применение стохастических моделей позволяет получать прогнозные оценки, в которых диапазон вероятных траекторий прямо отражает ту меру неопределенности, которая фактически присуща демографическому развитию мегаполиса.

Санкт-Петербург по своим демографическим характеристикам представляет собой показательный пример российского мегаполиса. На протяжении всей своей истории с момента его основания в 1703 г. город выступал в качестве барометра для демографических изменений, происходящих в российском обществе [Сафарова и др., 2015]. В истории города отчетливо прослеживаются периоды глубокого спада численности населения, вплоть до состояний демографической «клинической смерти» (потери Первой мировой войны, революций, голода и неустроенности, потери Великой Отечественной войны и блокады, кризис и депопуляция в начале 1990-х годов), и последующие этапы восстановления режима демографического воспроизводства, в значительной степени обеспеченные миграционным притоком. Последовательность этих фаз сформировала возрастную структуру с выраженными когортными волнами и высокой чувствительностью к изменениям миграционных потоков. В такой ситуации даже умеренные изменения уровней рождаемости, смертности или миграции приводят к заметным расхождениям в долгосрочных оценках численности и возрастной структуры населения. Эти изменения позволяют уловить

differing prediction intervals (e.g. Lee 1974, Cohen 1986, Sanderson 1995). This is also confirmed by our sensitivity tests. Therefore, we have most confidence in our short-term intervals (up to roughly 15–20 years ahead; at least for the TFR and the age structure), because these agree rather well with independently observed errors in historical forecasts. But our long run results must be interpreted with caution. We remind the reader of Joel Cohen's warning: "Uncertainty attaches not only to the point forecasts of future population, but also to the estimates of those forecasts uncertainty"» [Cohen 1986].

² Всероссийская перепись населения 2020 года.

³ Всероссийская перепись населения 2020 г. проведена по состоянию на 1 октября 2021 г.

⁴ Данные доклада «Социально-экономическое положение Санкт-Петербурга в январе–октябре 2021 года».

вероятностная когортно-компонентная модель, прослеживающая переходы когорт во времени и задающая рождаемость, смертность и миграцию как стохастические компоненты.

Цель настоящего исследования состоит в методологической адаптации вероятностного когортно-компонентного подхода, предложенного в ряде работ [Vanella, Deschermeyer, 2020; Носова, 2019; Yu et al., 2023; Cameron, Poot, 2011], к условиям открытой демографической системы крупного мегаполиса на примере Санкт-Петербурга. Задача работы заключается в том, чтобы на основе модифицированной когортно-компонентной схемы со стохастической параметризацией рождаемости, смертности и миграции получить распределения возможных траекторий численности и возрастной структуры населения Санкт-Петербурга и количественно оценить вклад основных компонент демографического движения в формирование медианных и крайних сценариев.

В мировой и отечественной практике основу демографических прогнозов до сих пор составляет детерминированный когортно-компонентный метод. На национальном уровне он работает надежно: сглаженная миграция и умеренные колебания рождаемости и смертности позволяют получить один согласованный сценарий будущей численности. В мегаполисе такая картина ломается: миграция становится ключевым фактором, определяющим динамику численности населения, при этом масштаб ее колебаний сопоставим с самим трендом, а фиксация единственного миграционного варианта представляет собой жесткое и малореалистичное допущение.

Поворот к вероятностным постановкам прогноза начался со страновых прогнозов ООН и постепенно сместился к региональному и городскому уровням. Сохраняя когортно-компонентный каркас, эти работы трактуют составляющие демографического прогноза не как заданные линии, а как стохастические процессы, допускающие несколько правдоподобных продолжений. Для крупного города обоснование метода прогнозирования численности и состава населения – это не формальный вопрос, а условие адекватности: без явного описания разброса миграционного режима население мегаполиса прогнозируется с заданной точностью.

Различие между традиционным и вероятностным когортно-компонентным подходом сводится к тому, что считать прогнозом: одну расчетную траекторию или распределение возможных траекторий. В первом случае неопределенность остается «по умолчанию» и проговаривается лишь в словесных сценариях. Во втором рождаемость, смертность и миграция задаются как случайные величины, которые колеблются вокруг ожидаемых уровней; такая модель многократно проигрывается в режиме Монте-Карло, а результатом становится медианная траектория с интервальными оценками, которые количественно описывают поле возможных будущих значений (численности населения или какой-то его составляющей).

2. Особенности и реализация модели

В основе прогноза лежит вероятностная когортно-компонентная модель с годовым шагом и разбиением населения по полу и пятилетним возрастным группам. Для возрастов $a > 0$ численность населения задается стандартным демографическим балансом:

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

где $P_{a,s,t}$ – численность населения пола s в группе a в начале года t ; $S_{a,s,t}$ – вероятность дожития; $M_{a,s,t}$ – сальдо миграции. Для группы 0–4 лет (новорожденные) используется рождаемость, рассчитываемая по формуле:

$$(2) \quad B_t = \sum_{a \in A_f} P_{a,t}^{(f)} ASFR_{a,t} \frac{5}{1000}, \quad P_{0,s,t+1} = B_t \cdot p_s,$$

где $ASFR_{a,t}$ – возрастные коэффициенты рождаемости; p_s – доли мальчиков и девочек при рождении, задаваемые через стабильное соотношение полов $SRB = 1,06$ (106 мальчиков на 100 девочек). Исходная возрастно-половая структура населения задается для базового года t_0 как P_{a,s,t_0} ; далее в каждом сценарии она развивается под действием стохастически заданных режимов рождаемости, смертности и миграции.

Режимы компонент задаются в два шага. На первом шаге исторические возрастные профили рождаемости, смертности и миграции за период наблюдений представляются в виде матриц «возраст × год» отдельно по полу. В логарифмической шкале каждый профиль аппроксимируется суммой нескольких главных компонент:

$$(3) \quad \ln \lambda_{a,s,t} = \mu_{a,s} + \sum_{k=1}^K z_{k,s,t} v_{k,a,s},$$

где $\mu_{a,s}$ – средний возрастной профиль; $v_{k,a,s}$ – возрастной паттерн k -й главной компоненты; $z_{k,s,t}$ – ее временной коэффициент; a – возраст; t – год; s – пол.

Для всех трех компонент (рождаемость, смертность, миграция) в дальнейших расчетах используются $K = 2$ главные компоненты. Выбор $K = 2$ обеспечивает компромисс между точностью аппроксимации и устойчивостью оценивания: первая компонента, как правило, отражает общий уровень возрастного профиля, а вторая – его основную деформацию (изменение формы профиля по возрастам). Это позволяет существенно сократить число параметров при сохранении интерпретируемости и последующего стохастического моделирования временных коэффициентов $z_{k,s,t}$.

На втором шаге временная динамика временных коэффициентов главных компонент $z_{k,s,t}$ ($k = 1, 2$; s – пол) описывается авторегрессионной моделью первого порядка:

$$(4) \quad z_{k,s,t} = \varphi_k z_{k,s,t-1} + \eta_{k,s,t}, \quad \eta_{k,s,t} \sim N(0, \sigma_k^2),$$

т.е. стационарным процессом AR(1), эквивалентным модели ARIMA(1,0,0). В качестве альтернатив для временных коэффициентов $z_{k,s,t}$ рассматривались модели ARIMA(p,d,q) с отбором по информационным критериям AIC и BIC (Приложение, табл. П1). Сравнение выполнялось для каждой компоненты ($k = 1, 2$) и пола на полном периоде 1990–2023 гг., а также на сокращенных периодах калибровки (1995–2023, 2000–2023, 2005–2023 гг.). Хотя по AIC во многих случаях предпочтение получали спецификации с $k = 1$, выбор по-

рядков (p,q) и тип лучшей модели (AR/ARMA/ARIMA) заметно менялись при смене периода калибровки и различались между компонентами и полами. Поскольку результаты отбора ARIMA(p,d,q) по AIC/BIC зависели от периода калибровки и различались между компонентами и полами, в качестве базовой выбрана парсимоничная модель AR(1), обеспечивающая воспроизводимую генерацию траекторий в Монте-Карло. Принцип парсимонии как основание выбора базовой спецификации широко используется в демографическом прогнозировании; в классической модели Ли – Картера авторы прямо отмечают, что при лишь незначительном улучшении более сложной версии они предпочли более простую динамику временного индекса⁵ [Lee, Carter, 1992]. В данном исследовании AR(1) используется не как эвристический инструмент, а как формализованная стохастическая модель динамики коэффициентов РСА, интегрированная в вероятностный когортно-компонентный прогноз. Такая спецификация позволяет учитывать инерционность демографических режимов, не накладывая жестких детерминированных трендов, и соответствует подходу, реализованному в ряде вероятностных прогнозов для стран и регионов [Носова, 2019; Yu et al., 2023; Cameron, Poot, 2011]. На прогнозном интервале генерируются 10000 стохастических траекторий $z_{k,s,t}$. По каждой из них восстанавливаются наборы возрастнополовых коэффициентов $\lambda_{a,s,t}^{(d)}$, которые далее используются в когортно-компонентной симуляции.

Смертность задается стохастическим коэффициентом смертности $m_{a,s,t}$. Вероятность дожития определяется экспоненциальной функцией:

$$(5) \quad S_{a,s,t} = \exp\left[-\left(m_{a,s,t} + \varepsilon_{m,a,s,t}\right)\right],$$

где $\varepsilon_{m,a,s,t}$ – отклонение от ожидаемого уровня дожития для лиц в возрасте a лет в t -ом году мужчин/женщин. Для самой старшей возрастной группы (85 лет и старше) рассчитанные прогнозные значения вероятности дожития дополнительно корректируются: коэффициенты смертности умножаются на множитель $OLD_AGE_MORT_MULT=1,05$, что повышает смертность на 5% и предотвращает нереалистичный рост дожития в условиях высокой статистической неопределенности на этих возрастах [Лаврикова и др., 2016]. Одновременно вводится верхний предел допустимых значений смертности по возрасту (CAP_BY_AGE), что исключает появление экстремальных реализаций при Монте-Карло-симуляции. Эти ограничения «подрезают» хвосты распределения, сохраняя при этом согласованность средних уровней смертности с наблюдаемой динамикой.

Миграционная компонента имеет двухуровневую постановку. Во-первых, сальдо миграции представляется через возрастнополовые уровни относительной чистой миграции $r_{a,s,t}$:

$$(6) \quad M_{a,s,t}^{(d)} = P_{a,s,t}^{(d)} \left(r_{a,s,t}^{(d)} + \varepsilon_{R,a,s,t}^{(d)} \right).$$

Переход от моделирования абсолютных потоков к относительным величинам позволяет учитывать демографический «размер» когорт и реалистично воспроизводить их

⁵ Ориг.: «A similar model with an AR(1) term added was marginally superior, but we preferred the (0, 1, 0) model on grounds of parsimony».

возрастную селективность, что особенно важно на субнациональном/региональном уровне [Yu et al., 2023; Cameron, Poot, 2011]. Исторические уровни относительной чистой миграции сглаживаются с помощью экспоненциально взвешенного среднего с параметром $\lambda = 0,8$, что повышает вес последних лет и лучше отражает современную структуру миграции при сохранении связи со всем рядом данных. Подбор параметра лямбда показал, что принятое основание для прогноза данных (2014–2023 гг.) минимизирует среднюю абсолютную ошибку воспроизведения фактического возрастного профиля миграции и позволяет корректно учесть вклад трудоспособных возрастов и возвратных потоков в старших группах.

Рождаемость формируется на основе сгенерированных возрастных коэффициентов $ASFR_{a,t}^{(d)}$ и текущей численности женщин репродуктивного возраста. Число рождений в году t определяется как

$$(7) \quad B_t^{(d)} = \sum_{a \in A_f} ASFR_{a,t}^{(d)} \cdot P_{a,t}^{(f,d)} \frac{5}{1000},$$

где $B_t^{(d)}$ – число рождений в году t ; $ASFR_{a,t}^{(d)}$ – возрастные коэффициенты рождаемости в возрастной группе a в году t ; $P_{a,t}^{(f,d)}$ – численность женщин репродуктивного возраста в возрастной группе a в году t ; A_f – множество возрастных групп женщин репродуктивного возраста; $\frac{5}{1000}$ – коэффициент пересчета, учитывающий пятилетний интервал возрастных групп и нормировку коэффициентов рождаемости на 1000 женщин.

Новорожденные распределяются по полу с использованием стабильного соотношения полов при рождении $SRB = 1,06$, как указано выше:

$$(8) \quad P_{male} = \frac{SRB}{1 + SRB}, \quad P_{female} = \frac{1}{1 + SRB}.$$

После задания компонентных соотношений модель реализуется в виде алгоритма Монте-Карло-симуляции. Полный алгоритм Монте-Карло-симуляции включает последовательное применение трех операций для каждой возрастно-половой когорты и каждого года: дожитие и переход в следующую возрастную группу, добавление миграции, рождение новой когорты. Эти операции реализуются в матричной форме, что обеспечивает согласованность демографического баланса и вычислительную устойчивость при числе симуляций $N = 10000$. Для каждого года прогноза формируется распределение $p_{a,s,t}^{(d)}$, на его основе вычисляются медианные значения $\tilde{p}_{a,s,t}$ и квантильные границы (например, q_5 и q_{95}) для численности населения по полу и укрупненным возрастным группам:

$$(9) \quad P_t^{(10)} = q_{0,05} \left(\sum_{a,s} P_{a,s,t}^{(d)} \right),$$

$$(10) \quad \tilde{P}_t = q_{0,50} \left(\sum_{a,s} P_{a,s,t}^{(d)} \right),$$

$$(11) \quad P_t^{(90)} = q_{0,95} \left(\sum_{a,s} P_{a,s,t}^{(d)} \right),$$

где $P_t^{(10)}$, \tilde{P}_t , $P_t^{(90)}$ – соответственно нижняя (5%), медианная (50%) и верхняя (95%) оценки общей численности населения в году t ; $P_{a,s,t}^{(d)}$ – численность населения пола s в возрастной группе a в году t в рамках одной реализации модели; $q_\alpha(\cdot)$ – эмпирическая α -квантиль распределения, полученного по результатам Монте-Карло-симуляции.

Ключевым элементом реализации является калибровка доверительных интервалов на основе ретроспективного моделирования (backtesting) [Архангельский, Елизаров, 2016; Raftery et al., 2012]. 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-процентный или 90-процентный интервалы.

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

$$(12) \quad C_p(h) = \frac{1}{|O|} \sum_{o \in O} I \left\{ P_{o+h}^{fact} \in [P_{o+h}^{(p_L)}, P_{o+h}^{(p_U)}] \right\},$$

где $|O|$ – число стартовых лет; $I\{\cdot\}$ – индикатор попадания; p_L и p_U – нижняя и верхняя границы интервала. Если эмпирическое покрытие $C_p(h)$ существенно ниже целевого (например, $C_{0,90}(5) = 0,72$), интервальные оценки расширяются путем умножения половины доверительного интервала на множитель:

$$(13) \quad k_p(h) = \frac{P_{target}}{C_p(h)}, \quad p \in \{0,75, 0,90\},$$

В использовании этот множитель ограничен диапазоном значений [1,0; 3,5]. Например, для горизонта, $h = 5$ лет, этот множитель составит $k_{0,90}(5) \approx 1,25$, что указывает на необходимость увеличить полуинтервал на 25%. После применения таких множителей по всем горизонтам достигается эмпирическое покрытие, близкое к 0,9 для 90-процентных и к 0,75 для 75-процентных интервалов. Как показал опыт региональных прогнозов численности населения, полученные характеристики сопоставимы с результатами backtest-калибровки для региональных прогнозов в Новой Зеландии и в округах США [Raftery et al., 2012; Ševčíková, Raftery, 2021].

3. Реализация методологии когортно-компонентного стохастического прогноза численности населения (на примере Санкт-Петербурга)

В прогнозном горизонте Санкт-Петербург выступает уже не как город, стоящий на пороге очередного демографического перелома, а как мегаполис, для которого ключевым становится удержание достигнутого масштаба населения. Вероятностный прогноз показывает, вокруг каких уровней и в каких пределах может колебаться численность жителей города в ближайшие десятилетия.

В реализации представленной методологии в качестве базового года принят 2023 г.: исходная возрастно-половая структура P_{a,s,t_0} задается по среднегодовой численности населения Санкт-Петербурга за 2023 г. (в разрезе пола и 5-летних возрастных групп). Официальный прогноз Росстата базируется на численности на начало 2023 г., поэтому при последующем сопоставлении траекторий учитывается различие в исходной базе. Горизонт прогноза до 2045 г. выбран для сопоставимости с последней версией официального прогноза по Санкт-Петербургу, опубликованной Росстатом.

Эмпирическая база для построения модели включает демографические данные за 1990–2023 гг.⁶: возрастные коэффициенты рождаемости *ASFR* и возрастно-половые коэффициенты смертности *ASMR* по 5-летним возрастным группам, а также данные о прибывших⁷ и выбывших⁸ мигрантах в абсолютных значениях. Данные за весь период используются для восстановления возрастных профилей демографических компонент, оценки их временной динамики и стохастической структуры, а также для проведения ретроспективного тестирования (*backtesting*) с целью определения прогностических свойств модели. Для оценки эмпирической возрастно-половой структуры миграционных потоков и формирования базового возрастного профиля миграции в прогнозе дополнительно используется период 2014–2023 гг.

Это связано с тем, что для миграции в модели требуется устойчивый возрастно-половой профиль (структура потоков), а не только уровни. На позднем интервале 2014–2023 гг. возрастно-половые распределения мигрантов более сопоставимы по методике учета и полноте, поэтому именно он используется для восстановления базового профиля миграции в прогнозе. При этом уровни миграции и динамика временных коэффициентов оцениваются на полном окне 1990–2023; интервал 2014–2023 используется только для стабилизации возрастной формы базового профиля.

Обоснование выбора периода эмпирической базы (1990–2023) и проверка чувствительности к альтернативным окнам. Выбор эмпирической базы для калибровки вероятностной модели определяет не только параметры динамики, но и масштаб оцениваемой неопределенности: стохастическая дисперсия временных коэффициентов главных компонент и, следовательно, ширина прогностических интервалов выводятся из

⁶ В условиях недостатка демографических данных за отдельные годы часть показателей за период 1990–2023 гг. была реконструирована и восстановлена автором методом вторичной группировки данных.

⁷ По данным ЕМИСС «Число прибывших по полу, возрасту и потокам передвижения» (<https://www.fedstat.ru/indicator/58613>)

⁸ По данным ЕМИСС «Число выбывших по полу, возрасту и потокам передвижения» (<https://www.fedstat.ru/indicator/58614>)

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.

В качестве базового принят интервал 1990–2023 гг., поскольку он обеспечивает полный охват качественно различных режимов демографической динамики. Для проверки устойчивости выбора исторического окна выполнена повторная перекалибровка модели на сокращенных периодах (1995–2023; 2000–2023; 2005–2023) с повторением PCA-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. Полученные прогнозные распределения действительно чувствительны к выбору периода калибровки: при сокращении базы меняются параметры стохастической динамики и, как следствие, ширина прогнозных интервалов. Именно поэтому выбор базового интервала обоснован через эмпирическую калибровку предиктивных интервалов (скользящий *backtesting*). Наиболее заметные расхождения проявляются в оценке неопределенности (ширине интервалов), тогда как медианные траектории меняются умеренно. Схожая зависимость устойчивости прогнозных оценок от длины базового периода отмечается и в классической постановке Ли – Картера: при достаточно длинной базе прогнозы устойчивы к сокращению окна, тогда как при слишком короткой базе проявляется нестабильность⁹ [Lee, Carter, 1992].

Сопоставление окон калибровки показывает, что переход к «посткризисному» периоду 2000–2023 гг. действительно меняет структуру данных: для рождаемости возрастает доля дисперсии, объясняемая первой компонентой (PC1=98,78%), тогда как для миграции, напротив, снижается вклад PC1 и возрастает вклад PC2, т.е. форма возрастного профиля становится более «двухкомпонентной» (см. табл. П2 Приложения).

Однако при переходе к окну 2000–2023 гг. сокращается наблюдаемая вариативность временных коэффициентов $z_{k,s,t}$, поэтому предиктивные интервалы становятся уже и возрастают риски недооценки неопределенности. Это проявляется при ретроспективной проверке (скользящий *backtesting*): эмпирическое покрытие некалиброванных интервалов заметно ниже номинального и в коротком окне ухудшается сильнее. Это подтверждается скользящим *backtesting* прогностических интервалов на горизонте до 5 лет: при калибровке на окне 2000–2023 гг. среднее эмпирическое покрытие составило 0,20 для 90-процентного интервала и 0,16 для 75-процентного интервала при номинальных уровнях 0,90 и 0,75 соответственно, т.е. интервалы оказываются чрезмерно узкими и не отражают масштаб возможных макродемографических шоков. При расширении базы до 1990–2023 гг. показатели покрытия возрастают, а требуемая корректировка калибровочных коэффициентов по горизонту оказывается существенно менее жесткой, что указывает на более реалистичную оценку стохастической неопределенности в долгосрочном прогнозе.

Отдельно следует подчеркнуть, что локальное преимущество моделей с интегрированием ($d = 1$) по AIC/BIC чаще проявляется при включении переходного периода 1990-х годов и не является устойчивым для всех компонент и подвыборок; при исключе-

⁹ Ориг.: «forecasts are insensitive to reductions in the length of the base period from 90 to 30 years» и «some instability appears for base periods of 10 or 20 years».

нии начала 1990-х результаты выбора d становятся менее однозначными. С учетом отсутствия систематического выигрыша по информационным критериям и критической роли корректной оценки неопределенности в вероятностном прогнозе базовое окно калибровки фиксируется на 1990–2023 гг., а результаты для сокращенных интервалов используются как анализ чувствительности и иллюстрация компромисса между структурной «чистой» посткризисной динамикой и реалистичностью оценки неопределенности.

Для эмпирической проверки достаточности принятого в модели числа компонент ($K = 2$) оценена доля дисперсии возрастных профилей, объясняемая первой и второй главными компонентами (PC1 и PC2), отдельно по полу и демографическим компонентам. Расчеты выполнены по рядам возрастных коэффициентов рождаемости (ASFR), смертности (ASMR) и возрастно-половых коэффициентов чистой миграции, восстановленных по данным о прибывших и выбывших, за 1990–2023 гг. Как показано в табл. 1, первые две компоненты совместно объясняют 81,6–98,7% вариации возрастных профилей, что подтверждает достаточность выбора $K = 2$ для последующего моделирования временных коэффициентов $z_{k,s,t}$.

Таблица 1.

**Доля первых двух главных компонент (PCA-AR(1))
в совокупной объясненной дисперсии**

Компонента	PC1, %	PC2, %
Смертность (F)	73,88	12,79
Смертность (M)	77,63	11,71
Миграция (F)	74,05	10,36
Миграция (M)	66,83	14,77
Рождаемость (F)	87,95	10,79

Источник: расчеты автора.

Дополнительно в модели учитывается возможный уровневый сдвиг миграции в будущем. По итогам ВПН-2020, недоучет численности населения Санкт-Петербурга по сравнению с текущим учетом оценивается примерно в 230 тыс. человек¹⁰. Если трактовать этот разрыв как накопленный за межпереписной период результат систематического занижения миграционного прироста, то он соответствует «скрытому» чистому притоку порядка 20–25 тыс. человек в год. Однако переносить весь этот объем в прогноз в виде ежегодной добавки некорректно: часть соответствующих когорт сформировалась на ранних этапах межпереписного периода и к моменту новой переписи уже частично выбыла из состава населения города, кроме того, недоучет имеет выраженную возрастную специфику и не может считаться равномерным по всем возрастно-половым группам.

Анализ возрастно-полового профиля официального миграционного сальдо показывает, что основной дополнительный приток приходится на молодые когорты. В 2023 г.

¹⁰ Для Санкт-Петербурга сопоставление итогов Переписи-2020 (5,602 млн человек на 01.10.2021 г.) и оценок текущего учета (5,3715 млн человек на 01.10.2021 г.).

наибольшее положительное сальдо наблюдалось в возрасте 15–19 лет (+8,576 тыс. человек), тогда как в группе 20–24 года фиксировалось отрицательное значение (–6,648 тыс. человек). Такая конфигурация отражает особенности статистического учета студенческой миграции: въезд сопровождается оформлением временной регистрации в общежитиях, тогда как по ее окончании значительная часть молодежи в возрасте 21–22 года, оставшаяся в городе, не переоформляет регистрацию по месту фактического проживания. В результате официальная статистика завышает масштабы выбытия в этих возрастах и занижает устойчивый миграционный прирост молодого населения.

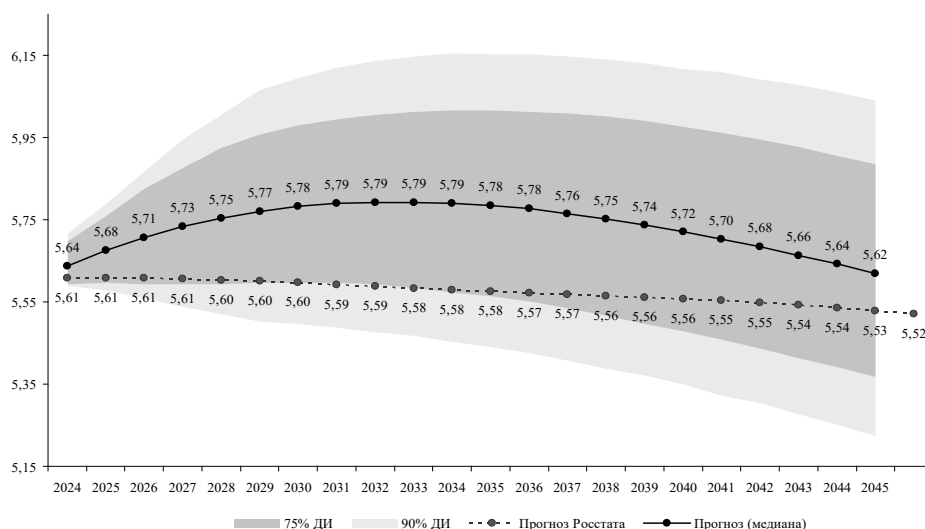
В модели этот структурный недоучет задается в виде дополнительно специфицируемого матричного потока $MIG_NET_BUMP = 8000$. Агрегированный по возрастно-половым группам эффект данного потока далее обозначается как *bump*-компонента миграционного баланса. Это примерно треть от оценочного годового скрытого притока, что соответствует предположению, что часть межпереписного разрыва связана с устойчивым закреплением мигрантов в городе и должна быть учтена при формировании базового уровня. Ежегодное сальдо миграции увеличивается на эту величину и распределяется по возрастно-половым группам согласно калиброванному треугольному профилю, концентрированному в возрастах 20–44 года с максимумом в группе 25–34 года. Подбор параметров показал, что при таком значении *bump*-компоненты миграционного баланса кумулятивный вклад дополнительной миграции согласуется как с переписным недоучетом численности, так и с наблюдаемой структурой трудовой миграции и не приводит к избыточной концентрации прироста в узком возрастном интервале.

4. Обсуждение результатов

Реализация методологии прогноза по данным Санкт-Петербурга показала, что в соответствии с медианной траекторией прогноза население мегаполиса в первой половине рассматриваемого периода будет продолжать умеренно расти, достигнув максимального уровня 5,792 млн человек в 2033 г., при 90-процентном доверительном интервале 5,468–6,155 млн человек, после чего темп прироста замедлится и сменится плавным спадом.

К 2045 г. медианная оценка численности составит 5,62 млн человек, что немного превышает уровень 2023 г. (5,598 млн человек). Центральный сценарий фактически фиксирует сохранение нынешнего масштаба населения с умеренными колебаниями, без резких взлетов и спадов. Разброс вероятных траекторий также невелик: 90-процентный доверительный интервал лежит в диапазоне $\pm 0,8$ млн человек (рис. 1). Понятно, что по мере удаления горизонта прогноза доверительный интервал несколько возрастает (до 0,82 млн человек), тем самым повышается неопределенность прогноза.

Медианный прогноз свидетельствует о переходе Санкт-Петербурга в фазу демографической стагнации. Рост в начале горизонта носит инерционный характер и не сопровождается восстановлением естественного воспроизводства. Эхо многочисленных когорт 2000-х годов проявляется лишь во временном смягчении снижения рождаемости по мере их вхождения в репродуктивные возраста, однако на фоне старения населения рост смертности доминирует, и естественный прирост становится устойчиво отрицательным, что со временем усиливается.



Источник: расчеты автора. Прогноз Росстата: «Предположительная численность населения Санкт-Петербурга до 2046». Стат. бюллетень. СПб.: Петростат, 2023. С. 6–7.

Рис. 1. Прогноз численности населения Санкт-Петербурга до 2045 г.: медиана, доверительные интервалы

В этих условиях миграция выполняет роль главного «амортизатора», компенсируя естественную убыль и удерживая численность населения примерно на достигнутом уровне. В медианном варианте чистая миграция остается положительной на всем горизонте, но постепенно снижается: от 34 тыс. человек в 2030 г. до 23 тыс. к 2045 г. Поэтому в первой половине горизонта миграционный прирост еще перекрывает естественную убыль и обеспечивает умеренный рост численности, но по мере ослабления миграции и углубления естественной убыли демографический баланс становится отрицательным, и траектория разворачивается к мягкому снижению.

Сопоставление результатов стохастического прогноза с официальными оценками выполнено на основе демографического прогноза Росстата, который был рассчитан детерминированным когортно-компонентным методом. В официальной методологии прогноз формируется от базовой численности населения на начало прогнозного периода (в актуальной версии – на начало 2023 г.) с учетом итогов ВПН-2020 и реализуется поэтапно с использованием процедуры «передвижки по возрастам». Сценарные условия задаются на уровне агрегированных демографических переменных: для рождаемости – суммарного коэффициента рождаемости и среднего возраста матери при рождении ребенка; для смертности – ожидаемой продолжительности жизни при рождении и коэффициента младенческой смертности; для миграции – абсолютных чисел прибывших и выбывших. Указанные сценарные переменные преобразуются в возрастно-половые коэффициенты рождаемости, смертности и миграции, после чего выполняется годовой расчет численности населения по полу и возрасту для каждого субъекта Российской Федерации. Далее результаты региональных расчетов балансируются с прогнозными итогами по России в целом.

До 2035 г. рассчитанная для Санкт-Петербурга траектория Росстата проходит в нижней части 75-процентных доверительных интервалов вероятностной модели; тем самым официальный сценарий задает более консервативную динамику. К концу периода различия становятся более умеренными: к 2045 г. Росстат ожидает 5,52 млн жителей, тогда как медианная оценка модели составляет 5,62 млн человек, так что расхождение составляет менее 2% (порядка 100 тыс. человек) и на всем горизонте данные официального прогноза полностью укладываются в 90-процентные доверительные интервалы. Наблюдаемый сдвиг траекторий в значительной степени обусловлен методологическими различиями в начальном уровне численности населения. В вероятностной модели в качестве базовой точки использовалась среднегодовая численность населения Санкт-Петербурга за 2023 г. по данным Петростата (5,5989 млн человек), тогда как официальный прогноз опирается на оценку на начало года (5,60004 млн человек). Это различие, вроде бы незначительное по величине, задает систематическое смещение траекторий.

Миграционная компонента моделировалась более детально: с использованием экспоненциально взвешенного сглаживания последних наблюдений, стохастического задания возрастно-половых относительных величин и дополнительной проверки ретроспективного расчета *birth*-эффекта в возрастах 20–44 лет, отражающего устойчивую притягательность города для молодежи и занятых. В результате стохастическая модель фиксирует несколько более устойчивое положительное миграционное сальдо по сравнению с консервативными допущениями Росстата. Наконец, при формировании исходной базы за 1990–2023 гг. для части показателей потребовались вторичная группировка и реконструкция из-за неполной сопоставимости детализированных статистических публикаций по Санкт-Петербургу в разные годы, что неизбежно порождает некоторые расхождения с официальными оценками¹¹.

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 гипотеза¹². По итоговой численности

¹¹ Под «дефицитом информации» понимается не отсутствие данных Росстата, а ограниченная сопоставимость и детализация показателей на длинном интервале при единой возрастной сетке, необходимой для когортно-компонентной модели. В частности, возрастно-половые коэффициенты смертности по 5-летним возрастным группам для Санкт-Петербурга в сопоставимом виде доступны начиная с 2011 г.; для 1990–2010 гг. использовались опубликованные укрупненные возрастные распределения и процедуры вторичной группировки/восстановления до 5-летней сетки. Для миграции детализированные потоки прибывших и выбывших по возрасту и полу формируются сопоставимо с 2003 г.; при этом в 2014 г. меняется возрастная классификация, поэтому данные 2003–2013 гг. были приведены к сетке после 2014 г. методом пропорционального перераспределения на основе усредненных долей 2014–2023 гг.; дополнительно возрастная группа 80+ согласовывалась с сеткой других компонент путем выделения подгрупп 80–84 и 85+.

¹² В стационарной авторегрессионной модели первого порядка AR(1) одношаговый прогноз после базового года в значительной мере определяется оцененными безусловным средним процессом и коэффициентом инерции. При $|\phi| < 1$ возникает механический эффект *mean reversion*: условное ожидание следующего значения $\mathbb{E}(x_{t+1}|x_t) = \mu + \phi(x_t - \mu)$, поэтому прогнозируемое значение x_{t+1} смещается от последнего наблюдения x_t в сторону безусловного среднего μ . Если ряд неод-

населения средний вариант Росстата на всем прогнозном горизонте остается внутри доверительного коридора вероятностной модели.

По СКР¹³ к 2030 г. различие минимально: медиана авторского прогноза составляет 1,325 против 1,334 в среднем варианте Росстата. Далее контуры расходятся: Росстат задает 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 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 материнству¹⁴. 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 остается 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 ее компенсирующая роль по отношению к естественной убыли сохраняется дольше; различие уровней связано и с иной калибровкой базовой траектории, и с различием способа задания миграционных предпосылок¹⁵.

В смертности контраст проявляется одновременно в траекториях ОПЖ¹⁶ 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; начальный рост не переходит в устойчивую траекторию улучшения. В среднем варианте Росстата заложено более выражен-

нороден 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.

¹³ 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.

¹⁴ Под многочисленными когортами понимаются поколения, сформированные на волне подъема 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 объем рождений.

¹⁵ 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 возрастнополовые коэффициенты и включаются в расчет по процедуре «передвижки по возрастам»; сценарии уточняются раз в два года.

¹⁶ 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.

ное 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 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. 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.

Поскольку средний вариант прогноза Росстата основан на экспертно заданных сценарных переменных демографических процессов, далее базовая вероятностная траектория дополняется авторским сценарным анализом.

Сценарные оценки. Важным дополнением стохастической модели выступает сценарный прогноз, позволяющий отразить возможные структурные сдвиги в рождаемости, смертности и миграции. Сформированы четыре варианта сценариев – от условно благоприятного до условно неблагоприятного, задающие реалистичный диапазон траекторий демографического развития мегаполиса (см. табл. 2).

Таблица 2.

Сценарные поправки к возрастным коэффициентам рождаемости и смертности и дополнительное годовое чистое сальдо миграции

Сценарий	Изменение возрастных коэффициентов рождаемости, %	Изменение возрастных коэффициентов смертности, %	Дополнительное годовое чистое сальдо миграции, тыс. человек
C1	+5	-5	+35
C2	+3	-3	+25
C3	-2	+2	+15
C4	-5	+5	0

Примечание: проценты отражают сценарные поправки к возрастным коэффициентам рождаемости и смертности. Итоговые числа родившихся и умерших сформированы в ходе когортно-компонентного расчета с учетом возрастной структуры населения. Поправки ±% применяются к базовой стохастической траектории на уровне возрастных коэффициентов: в каждой симуляции возрастные коэффициенты рождаемости (ASFR) и смертности (ASMR) масштабируются мультипликативно (fert_mult, mort_mult). Поправка к миграции задается как ежегодная аддитивная добавка к чистому миграционному приросту (mig_bump), распределяемая по возрастно-половым группам с концентрацией в 20–44 года по заданным весам.

Сценарный блок позволяет оценить, как изменение внешних и внутренних условий отразится на динамике населения города, поскольку в ситуации структурных сдвигов экстраполяция прошлых трендов оказывается недостаточной и, как отмечает Е.М. Андреев, «невозможен ни традиционный трендовый, ни нормативно-трендовый подход; единственный путь – разработка набора прогнозных сценариев» [Андреев, 1999]. В качестве верхней границы рассматривается комплексный благоприятный сценарий (C1). Он предполагает

ет одновременное улучшение всех компонент демографической динамики: умеренный рост рождаемости и снижение смертности (по +5 и -5% к базовым коэффициентам) в сочетании с дополнительным миграционным притоком 35 тыс. человек ежегодно. Такой вариант соответствует ситуации, когда внутренние демографические меры дают ощутимый эффект, а экономическая и институциональная привлекательность города поддерживает высокий уровень притока. На противоположном полюсе находится сценарий демографической стагнации (С4): интенсивность рождаемости на 5% ниже базового уровня, интенсивность смертности на 5% выше, дополнительный миграционный прирост отсутствует, чистое сальдо миграции близко к нулю. По сути, это траектория мегаполиса без притока молодых мигрантов, при которой внутренние демографические резервы постепенно исчерпываются.

Два промежуточных варианта фиксируют асимметричные сочетания внутренних и внешних условий. Сценарий опоры на внутренние демографические резервы (С2) предполагает частичное улучшение интенсивности рождаемости и смертности ($\pm 3\%$ к базовому уровню) при более сдержанном дополнительном притоке (около 25 тыс. человек в год). Напротив, сценарий внешней миграционной компенсации (С3) соответствует умеренному ухудшению внутренних режимов (-2% к возрастным коэффициентам рождаемости и +2% к возрастнo-половым коэффициентам смертности), которое сглаживается за счет сохраняющегося внешнего миграционного ресурса - 15 тыс. человек ежегодно.

Величины корректировок заданы по ретроспективному диапазону наблюдавшихся отклонений и используются как ориентир масштаба возможных сдвигов, а не как утверждение о буквальном повторении прошлых колебаний. Тем самым сценарии интерпретируются как проверка чувствительности базовой траектории к изменениям условий, которые не извлекаются из инерции временных рядов. На практике такие сдвиги для мегаполиса могут быть связаны с изменением экономической конъюнктуры и стабильности рубля, политического контекста и миграционного законодательства, а также с тем, насколько предсказуемы институциональные условия жизни и работы для приезжающих и остающихся в городе.

Сравнение сценарных траекторий показывает, что в краткосрочной перспективе различия между ними невелики, однако со временем они нарастают, что приводит к формированию множества треков (рис. 2).

Уже к середине прогнозного горизонта, в окрестностях 2035 г., сценарные кривые начинают расходиться не как небольшие поправки единой траектории, а как разные варианты демографического прогноза. Если в 2025 г. разница между крайними сценариями составляет лишь несколько десятков тысяч человек (что закономерно: ближайшие годы во многом «защиты» в уже сложившейся возрастной структуре), то к 2035 г. разрыв достигает 600-700 тыс. человек. В 2045 г. веер вариантов раскрывается полностью: при комплексе благоприятных предпосылок численность населения выходит на уровень 6,4 млн человек, при сценарии демографической стагнации снижается до 5,32 млн, т.е. диапазон вариации превышает 1,1 млн. Промежуточные сценарии показывают более узкий коридор, при несколько более высокой стартовой численности - от 5,75 до 6,1 млн жителей. Базовая медианная оценка (5,62 млн) лежит ниже середины этого коридора и ближе к умеренно неблагоприятной траектории, что подчеркивает осторожность центрального прогноза, без закладывания максимально благоприятных допущений по всем компонентам демографического баланса.

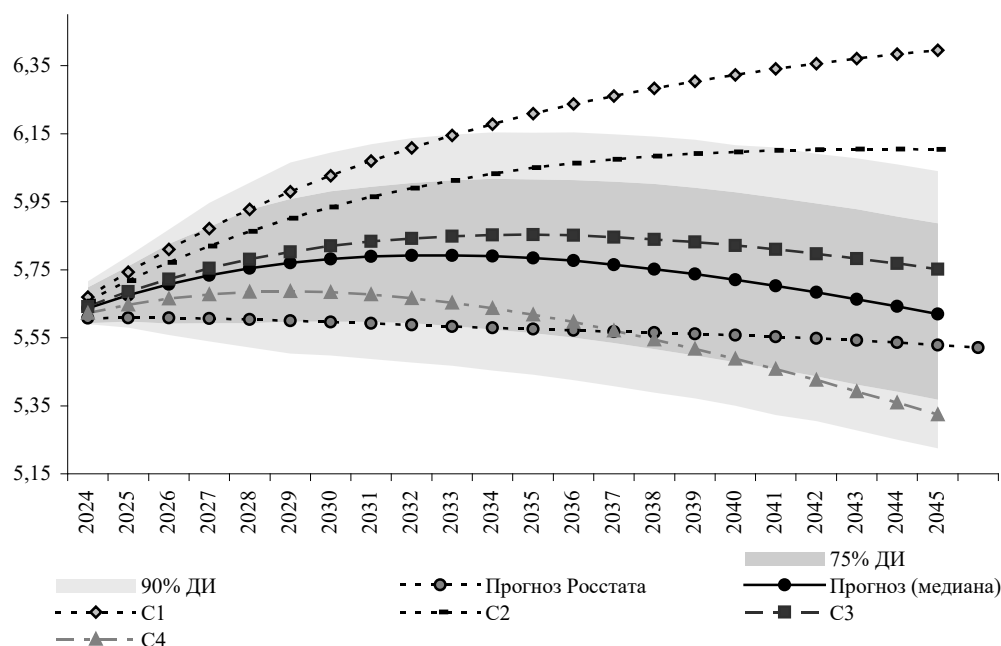


Рис. 2. Сопоставление сценарных прогнозов численности населения Санкт-Петербурга с оценками Росстата, 2024–2045 гг.

При этом сценарный анализ довольно ясно показывает, какая именно составляющая «держит» долгосрочную численность населения. Изменения в рождаемости и смертности в реалистичных пределах перерисовывают возрастную пирамиду, но сами по себе не меняют порядок величины общей численности; главным фактором, раздвигающим сценарные траектории к 2040-м годам, становится накопленный миграционный эффект. По сути, ответ на вопрос, будет ли через двадцать лет в Петербурге около 5,3 или свыше 6 млн жителей, определяется состоянием миграционного канала. При этом в прогнозном горизонте предполагается сохранение сложившейся структуры чистого миграционного прироста: 68–70% его объема формируется межрегиональной миграцией и около 28–30% – международной¹⁷. При устойчивом чистом притоке порядка 30–40 тыс. человек в год город не только компенсирует естественную убыль, но и сможет закрепить численность населения, превысив 6 млн жителей. При почти закрытом миграционном «клапане» численность медленно смещается к 5,3 млн, к нижней границе доверительного интервала базового прогноза. В этом смысле миграция в мегаполисе задает наиболее «быстродействующий» регулятор масштаба населения, тогда как рождаемость и смертность – его базовую де-

¹⁷ Рассчитано автором методом экстраполяции структуры чистого миграционного сальдо по данным текущего учета миграции Санкт-Петербурга за 2017–2023 гг. (соотношение межрегиональной и международной составляющих в суммарном сальдо за период), ЕМИСС. (<https://www.fedstat.ru/indicator/58615>)

мографическую «инерцию»: они формируют будущие поколения, а значит, определяют траекторию уже за пределами текущего горизонта и масштаб структурного старения.

Полученные результаты показывают, при каких сочетаниях рождаемости, смертности и миграции меняется масштаб населения мегаполиса в долгосрочном горизонте. Стохастическая модель задает базовый коридор неопределенности, в рамках которого лежат как медианный прогноз, полученный стохастическим когортно-компонентным прогнозом, так и официальная траектория Росстата. Сценарные кривые частично «раздвигают» этот коридор: неблагоприятный вариант опускается к нижней границе 90-процентного интервала ($\approx 5,3$ млн человек к 2045 г.), а два сценария с усиленным миграционным притоком выводят численность выше верхней кромки базового интервала, достигая 6,1–6,4 млн. Тем самым сценарный блок показывает, как именно будет меняться размер города при целенаправленном смещении демографических режимов.

Коридор между двумя «средними» вариантами на горизонте 2045 г. ограничен интервалом примерно 5,75–6,1 млн человек, и базовая оценка 5,62 млн располагается у его нижней границы. Иными словами, медианный прогноз можно интерпретировать как осторожно консервативный: он не закладывает ни резкого уменьшения численности населения, ни максимально возможной мобилизации миграционного ресурса, а описывает продолжение нынешней траектории с умеренным ослаблением миграционного притока.

Сравнение с официальным прогнозом Росстата подтверждает устойчивость выводов. Вероятностная постановка позволяет интерпретировать прогноз не как набор дискретных линий, а как распределение траекторий и их квантилей. Тем самым официальная траектория и сценарные варианты сопоставляются с базовой траекторией по положению внутри распределения, а не только по разнице уровней. Поскольку прогноз строится в когортно-компонентной форме, неопределенность переносится на возрастно-половой состав и производные показатели. Это делает возможным обсуждение не только «сколько жителей», но и рисков по структуре населения и по компонентному балансу.

До середины 2030-х годов кривая Росстата для Санкт-Петербурга идет вдоль нижней кромки доверительного интервала, затем постепенно сближается с медианной траекторией и к 2045 г. дает значение 5,52 млн против 5,62 млн на основе стохастической модели. Разрыв менее 2% и полностью укладывается в разброс, обусловленный альтернативными допущениями о миграции. По сути, оба подхода описывают близкую картину: отсутствие взрывного роста и нарастающее влияние старения населения на демографический баланс мегаполиса.

В этих условиях на первый план выходит не столько оценка «сколько жителей будет в Петербурге», сколько вопрос «каким будет их состав». А.А. Сафарова и Г.Л. Сафарова отмечают, что проблема старения для Санкт-Петербурга носит давний характер: уже с 1994 г. доля пожилых людей превысила долю детей в общей численности населения города [Сафарова и др., 2015]. В медианном варианте доля населения 65+ увеличится с 18,36% в 2024 г. до 23% к 2045 г., а численность ядра трудоспособных возрастов 20–64 сократится с 3,53 млн человек до 3,37 млн человек. Следовательно, миграция может «перекрыть» негативный эффект старения только в том случае, если она останется не просто положительной по объему, но и омолаживающей по структуре.

Меру необходимой миграционной компенсации задает простой демографический баланс: в 2045 г. медианное число рождений составит 41,57 тыс. человек, число умерших будет равно 95,00 тыс., естественная убыль достигнет 53,45 тыс. человек. При ме-

дианном чистом сальдо миграции 23,40 тыс. удержание численности на достигнутом «плато» к концу горизонта потребует чистого миграционного прироста 53,45 тыс. человек в год, т.е. фактически полного перекрытия естественной убыли. Однако в условиях старения решающее значение имеет не столько величина миграционного прироста, сколько то, какие когорты он пополняет: компенсирующий эффект может обеспечивать приток, сосредоточенный в молодых трудоспособных возрастах (прежде всего 20–39 лет), не уходящий в старшие когорты и без выраженного перекоса по полу. В этом случае миграция работает как демографический «регулятор»: она пополняет численность трудоспособных когорт и замедляет рост доли населения 65+ за счет притока молодых когорт, может положительно повлиять на рождаемость и снизить смертность. Поэтому главный итог прогноза можно сформулировать позитивно: при сохранении устойчивого омолаживающего миграционного притока Петербург способен удерживать и расширять масштаб населения и одновременно смягчать структурные последствия старения.

Представленный вероятностный когортно-компонентный подход показывает, что традиционная когортно-компонентная модель может быть существенно обогащена для условий открытых демографических систем. Сочетание PCA-AR(1)-моделирования компонент, задания миграции в виде возрастно-половых относительных величин с дрейфом и эмпирической калибровки доверительных интервалов позволяет формировать методический инструментарий демографического прогнозирования, особенно востребованный для крупных городов, где миграция играет ключевую роль в формировании численности и возрастной структуры населения. Вместе с тем предложенный подход применим не только к мегаполисам: он может служить методологическим ориентиром для совершенствования субнациональной практики демографического прогнозирования как крупнейших городов, так и регионов в России и других странах.

Приложение.

Таблица П1.

**Лучшая ARIMA(p,d,q) по критерию AIC
и приращение относительно ARIMA(1,0,0)**

Блок	Пол	k	Лучшая модель	ΔAIC
Период калибровки: 1990–2023 гг.				
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

Окончание табл. П1.

Блок	Пол	k	Лучшая модель	ΔAIC
3				
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 гг.				
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 гг.				
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.

Таблица П2.

**Сравнение доли объясненной дисперсии (PC1 и PC2)
при PCA-разложении возрастных профилей
на окнах 1990–2023 и 2000–2023 гг., %**

Компонента	PC1		PC2	
	1990–2023 гг.	2000–2023 гг.	1990–2023 гг.	2000–2023 гг.
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

* *

*

СПИСОК ЛИТЕРАТУРЫ

Андреев Е.М. Современный демографический кризис и прогнозы населения России // Мир России. Социология. Этнология. 1999. Т. 8. № 4. С. 175–186.

Архангельский В.Н., Елизаров В.В. Демографические прогнозы в современной России: анализ результатов и выбор гипотез: Научные труды. Институт народнохозяйственного прогнозирования РАН, 2016. 14. С. 524–545.

Лаврикова Ю.Г., Антипин И.А., Прядеин А.А., Суворова А.В. Прогноз развития крупнейшего города: конструирование инновационного будущего // Экономические и социальные перемены: факты, тенденции, прогноз. 2016. № 6(48). С. 214–235. <https://doi.org/10.15838/esc.2016.6.48.12>

Никитина С.Ю., Щербов С.Я. Вероятностный прогноз численности населения России // Вопросы статистики. 2007. № 7. С. 6–9.

Носова М.Г. Стохастическая модель для прогнозирования численности населения // Бюллетень науки и практики. 2019. Т. 5. № 9. С. 17–25. <https://doi.org/10.33619/2414-2948/46/02>

Сафарова А.А., Сафарова Г.Л., Косолапенко Н. Г., Арутюнов А.В. Демографические аспекты старения населения Санкт-Петербурга в конце XX – начале XXI в. Часть 1. Традиционные меры старения населения // Успехи геронтологии. 2015. Т. 28. № 4. С. 605–611.

Cameron M., Poot J. Lessons from Stochastic Small-Area Population Projections: The Case of Waikato Subgroups in New Zealand // Journal of Population Research. 2011. 28. P. 245–265. <https://doi.org/10.1007/s12546-011-9056-3>

Keilman N., Pham D., Hetland A. Why Population Forecasts Should Be Probabilistic: Illustrated by the Case of Norway // Demographic Research. 2002. 6. P. 409–454. <https://doi.org/10.4054/DemRes.2002.6.15>

Lee R.D., Carter L.R. Modeling and Forecasting U.S. Mortality // Journal of the American Statistical Association. 1992. 87(419). P. 659–671. <https://doi.org/10.1080/01621459.1992.10475265>

Raftery A.E., Li N., Ševčíková H., Gerland P., Heilig G.K. Bayesian Probabilistic Population Projections for All Countries. Proceedings of the National Academy of Sciences of the United States of America. 2012. 109(35). P. 13915–13921. <https://doi.org/10.1073/pnas.1211452109>

Ševčíková H., Raftery A. Probabilistic Projection of Subnational Life Expectancy // Journal of Official Statistics. 2021. 37. P. 591–610. <https://doi.org/10.2478/jos-2021-0027>

Vanella P., Deschermeier P. A Probabilistic Cohort-Component Model for Population Forecasting: The Case of Germany // Population Ageing. 2020. 13. P. 513–545. <https://doi.org/10.1007/s12062-019-09258-2>

Yu C.C., Ševčíková H., Raftery A.E., Curran S.R. Probabilistic County-Level Population Projections // Demography. 2023. 60(3). P. 915–937. <https://doi.org/10.1215/00703370-10772782>

Автор выражает искреннюю благодарность члену-корреспонденту РАН Елисейевой И.И. за помощь в обработке и интерпретации данных.

Population Forecasting in a Modern Metropolis (using St. Petersburg as a Case Study): Methodological Aspects

Daria Ryabchikova

Saint-Petersburg State University of Economics (UNECON).
30–32, bldg A, Griboyedov Canal Embankment, Saint Petersburg, 191023, Russian Federation.
E-mail: d.sinitzyna@yandex.ru

This article examines probabilistic cohort-component forecasting of the population of a megacity as an open demographic system, where migration determines the dynamics of the population size and age structure. In the methodology under consideration, fertility, mortality, and migration are modeled as stochastic age-sex processes: age profiles are reduced using PCA, the dynamics of time coefficients are described by AR(1) (equivalent ARIMA (1,0,0)), and the forecast is generated as a distribution of trajectories in a Monte Carlo simulation. The migration component is specified through relative age-sex rates with drift and an additional level adjustment (bump component) reflecting the possible underestimation of migration growth; interval estimates are validated by backtesting and calibrated against empirical coverage. All calculations are based on St. Petersburg data; the empirical base covers 1990–2023, and the forecast horizon extends to 2045, ensuring comparability with official Rosstat estimates. The median trajectory marks the city's transition to a demographic plateau: by 2045, the population will reach 5.62 million, with a 90% range of 5.225–6.041 million. Simultaneously, population aging accelerates (the share of residents aged 65+ reaches 23% by 2045), and the role of migration shifts from a source of growth to a mechanism for compensating for natural decline and stabilizing the age structure. Scenario forecasts show that the long-term fork in the trajectory is determined primarily by the cumulative migration effect (by 2045, from 5.32 million to 6.40 million in extreme cases). A comparison with the Rosstat forecast reveals moderate discrepancies; the official forecast trajectory falls within the confidence intervals of the probability model. The proposed toolkit can be used for population forecasting at the level of major cities and regions across the country.

Daria Ryabchikova – Postgraduate student.

Key words: stochastic cohort component model; Monte Carlo; PCA-AR(1); backtesting; migration; open demographic system; megacity; St. Petersburg, population aging.

JEL Classification: J11, J61, C53.

* *
*

References

Andreev E.M. (1999) The Contemporary Demographic Crisis and Population Forecasts for Russia. *Mir Rossii (Universe of Russia): Sociology, Ethnology*, 8, 4, pp. 175–186. (In Russ.)

Arkhangelsky V.N., Elizarov V.V. (2016) *Demographic Forecasts in Modern Russia: Analysis of Results and Choice of Hypotheses*. Scientific Works: Institute of Economic Forecasting, Russian Academy of Sciences, 14, pp. 524–545. (In Russ.)

Cameron M., Poot J. (2011) Lessons from Stochastic Small-Area Population Projections: The Case of Waikato Subgroups in New Zealand. *Journal of Population Research*, 28, pp. 245–265. <https://doi.org/10.1007/s12546-011-9056-3>

Keilman N., Pham D., Hetland A. (2002) Why Population Forecasts Should Be Probabilistic: Illustrated by the Case of Norway. *Demographic Research*, 6, pp. 409–454. <https://doi.org/10.4054/DemRes.2002.6.15>

Lavrikova Yu. G., Antipin I. A., Pryadein A. A., Suvorova A. V. (2016) Forecasting the Development of the Largest City: Designing an Innovative Future. *Economic and Social Changes: Facts, Trends, Forecast*, 6, 48, pp. 214–235. (In Russ.) <https://doi.org/10.15838/esc.2016.6.48.12>

Lee R.D., Carter L.R. (1992) Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association*, 87, 419, pp. 659–671. <https://doi.org/10.1080/01621459.1992.10475265>

Nikitina S.Yu., Shcherbov S.Ya. (2007) Probabilistic Forecast of the Population Size of Russia. *Vo-prosy Statistiki*, 7, pp. 6–9. (In Russ.)

Nosova M.G. (2019) Stochastic Model for Population Forecasting. *Bulletin of Science and Practice*, 5, 9, pp. 17–25. <https://doi.org/10.33619/2414-2948/46/02>

Raftery A.E., Li N., Ševčíková H., Gerland P., Heilig G.K. (2012) Bayesian Probabilistic Population Projections for All Countries. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 35, pp. 13915–13921. <https://doi.org/10.1073/pnas.1211452109>

Safarova A.A., Safarova G.L., Kosolapenko N.G., Arutyunov A.V. (2015) Demographic Aspects of Population Ageing in Saint Petersburg at the End of the 20th and the Beginning of the 21st Century. Part 1. Conventional Measures of Population Ageing. *Advances in Gerontology*, 28, 4, pp. 605–611. (In Russ.)

Ševčíková H., Raftery A. (2021) Probabilistic Projection of Subnational Life Expectancy. *Journal of Official Statistics*, 37, pp. 591–610. <https://doi.org/10.2478/jos-2021-0027>

Vanella P., Deschermeier P. (2020) A Probabilistic Cohort-Component Model for Population Forecasting: The Case of Germany. *Population Ageing*, 13, pp. 513–545. <https://doi.org/10.1007/s12062-019-09258-2>

Yu C.C., Ševčíková H., Raftery A.E., Curran S.R. (2023) Probabilistic County-Level Population Projections. *Demography*, 60, 3, pp. 915–937. <https://doi.org/10.1215/00703370-10772782>