



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

ПРОГНОЗЫ НАСЕЛЕНИЯ

Денисенко М.Б.,

Институт демографии имени А.Г.Вишневого НИУ ВШЭ



«Так как цель науки есть предвидение, то с того момента, когда оно становится невозможным, мы находимся вне пределов науки»

Эмиль Борель, французский математик

**«ЗНАТЬ, ЧТОБЫ ПРЕДВИДЕТЬ, ПРЕДВИДЕТЬ, ЧТОБЫ
ДЕЙСТВОВАТЬ»**

**Огюст Конт,
ОСНОВОПОЛОЖНИК СОЦИОЛОГИИ**

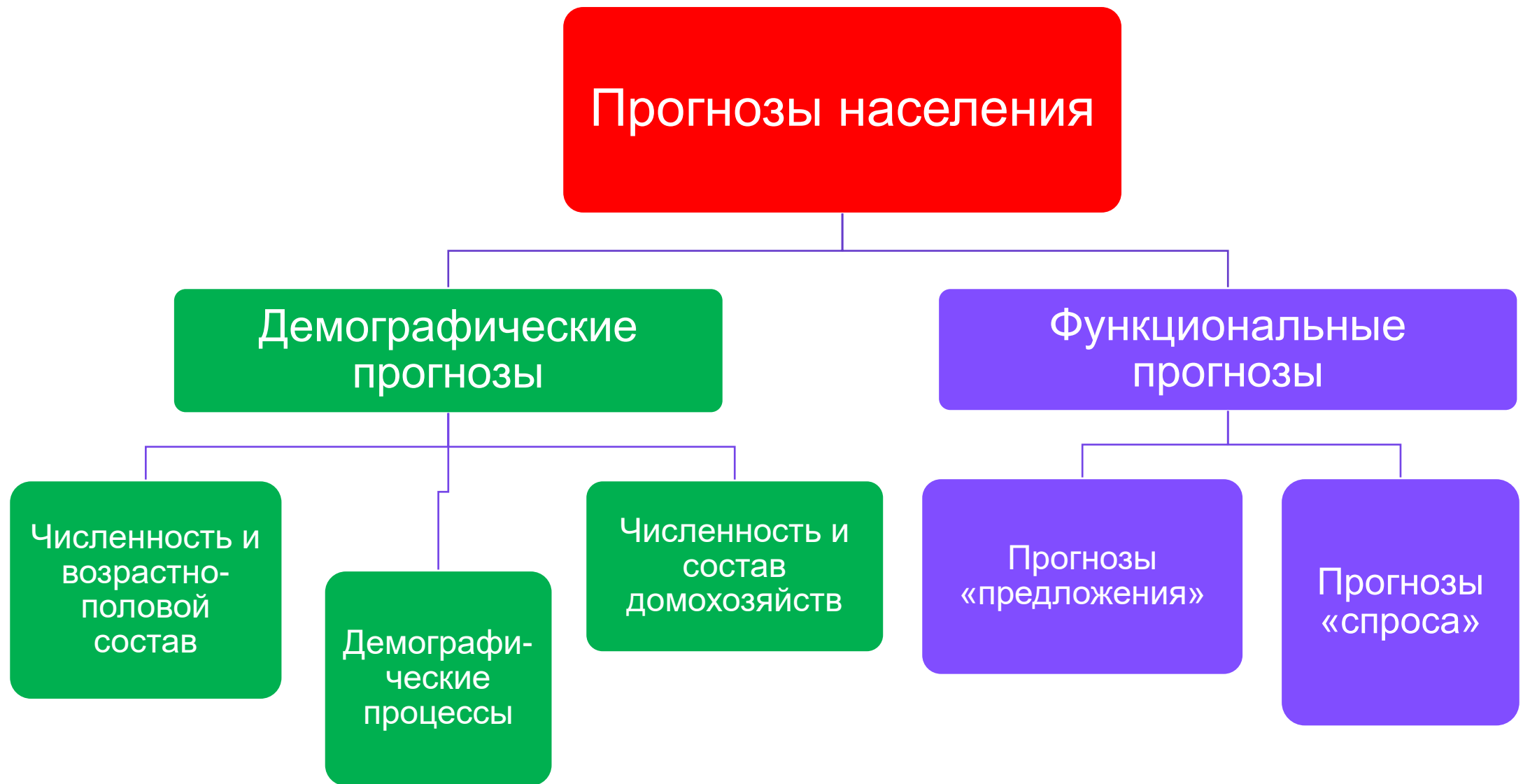
Темы

- Прогнозы населения: определения, классификации, история, точность прогнозов.
- Методы построения демографических прогнозов: население без структуры, когортно-компонентный метод, прогностические сценарии демографических процессов
- Функциональные прогнозы: подходы к построению, прогноз рабочей силы, прогноз уровня образования населения, агент-ориентированные прогнозы

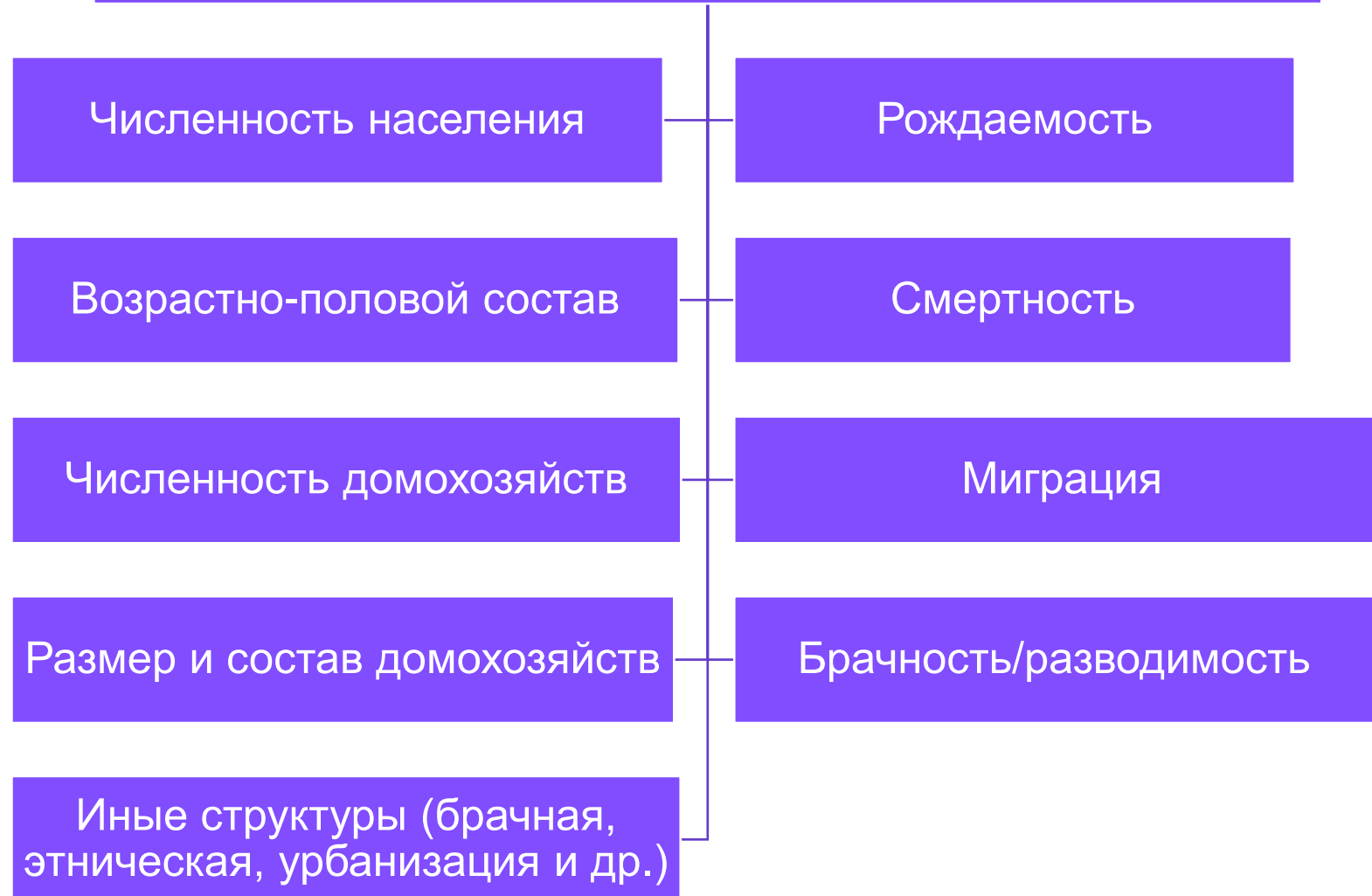
Прогноз (от греч. Prognosis) - предвидение, предсказание

Термин «прогноз», как предсказание состояния любого явления в будущем, пришел из медицины, где врачи делали «прогноз» об исходе болезни.

В науке термином «прогноз» обозначают вероятностное суждение о будущем, основанное на результатах специальных научных исследований.



Демографические прогнозы



Функциональные прогнозы

Функциональные группы

- совокупности людей, с которыми связана деятельность различных организаций, предприятий, отраслей производства и пр.

Функциональные прогнозы

- Прогнозы численности и состава функциональных групп называют функциональными прогнозами.

Для чего нужны демографические прогнозы? Функциональные прогнозы?

Оценка перспективной численности и состава населения, домохозяйств, параметров демографических процессов

Оценка численности и возрастно-полового состава отдельных функциональных групп

определение потребностей в продовольствии, энергии, жилье, в услугах системы образования, здравоохранения и т.д.

подготовка и обоснование прогнозов социально-экономического развития страны и ее регионов

разработка социально-экономической и демографической политики

Глобальные прогнозы изменений в геополитической ситуации

Прогнозы
развития

Функциональные
прогнозы

Демографические
прогнозы

Виды прогнозов

1. Реалистический прогноз. Цель: получить наиболее вероятные оценки будущего. **То, что нужно управленцам.**
2. Аналитический прогноз. Задача – изучить взаимосвязи между демографическими переменными. Например, **какой будет возрастная структура, если резко снизится рождаемость .**
3. Алармистские прогнозы или прогнозы-предостережения (вариант аналитического прогноза). Задача: **показать, чего следует ожидать в случае продолжения негативных тенденций развития .**

Термин «Прогноз» и «Оценки населения» в англоязычной литературе:

- **population forecast** – реалистический прогноз, наиболее вероятный, по мнению экспертов, прогноз
- **population projection** – прогнозы населения вообще (аналитические прогнозы, разные варианты прогнозов). Название отражает главный метод прогнозирования – проекцию на будущее прошлых тенденций
- **population estimates** – оценки численности и возрастно-полового состава населения, полученные с помощью уравнения демографического баланса и/или ретроспективного прогноза.

Прогнозы и время

Горизонт прогнозирования

Интервал между датой целевого (конечного) года прогноза и датой базового (начального) года прогноза

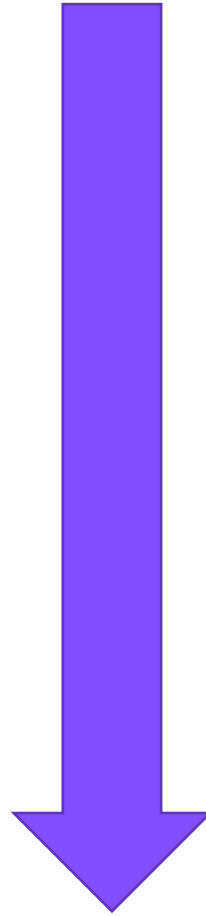
- Краткосрочные: до 10 лет;
- Среднесрочные: 10 - 25 лет;
- Долгосрочные (более 25 лет или прогноз за пределы «длины поколения»)

Шаг прогнозирования

- Интервал между датами (точками), на которые выполняется прогноз.
- Если задана возрастная структура, то шаг определяется длиной ее возрастного интервала.
- Результаты между опорными точками прогноза получают методом интерполяции

Географические объекты демографического прогнозирования

- Население мира и его регионов
- Население стран
- Население регионов страны и крупных городов (региональные прогнозы)
- Малые территории (низовые административно-территориальные и муниципальные образования)



**Трудности в прогнозировании
возрастают с уменьшением
численности населения
территории из-за уменьшения
числа демографических
событий и усиления роли
миграции в ее динамике**

Демографические прогнозы **МНОГОВАРИАНТНЫ**

Детерминистские (сценарные) прогнозы

- В каждый момент t демографическая функция принимает одно значение

На практике строят, как минимум, три варианта прогноза:

1. **нижний (худший, пессимистический);**
2. **средний (наиболее вероятный);**
3. **высокий (лучший, оптимистический).**

Варианты для процессов являются сценариями для прогноза численности и возрастного состава населения

Вероятностные прогнозы

Значения демографической функции – случайные величины, которые в каждый момент времени могут принимать разные значения из интервала.

Результаты прогноза – интервалы изменения прогнозируемой величины (интервальные прогнозы).

Средний (медиана) вариант – наиболее вероятный.

Отражает неопределённость изменений.

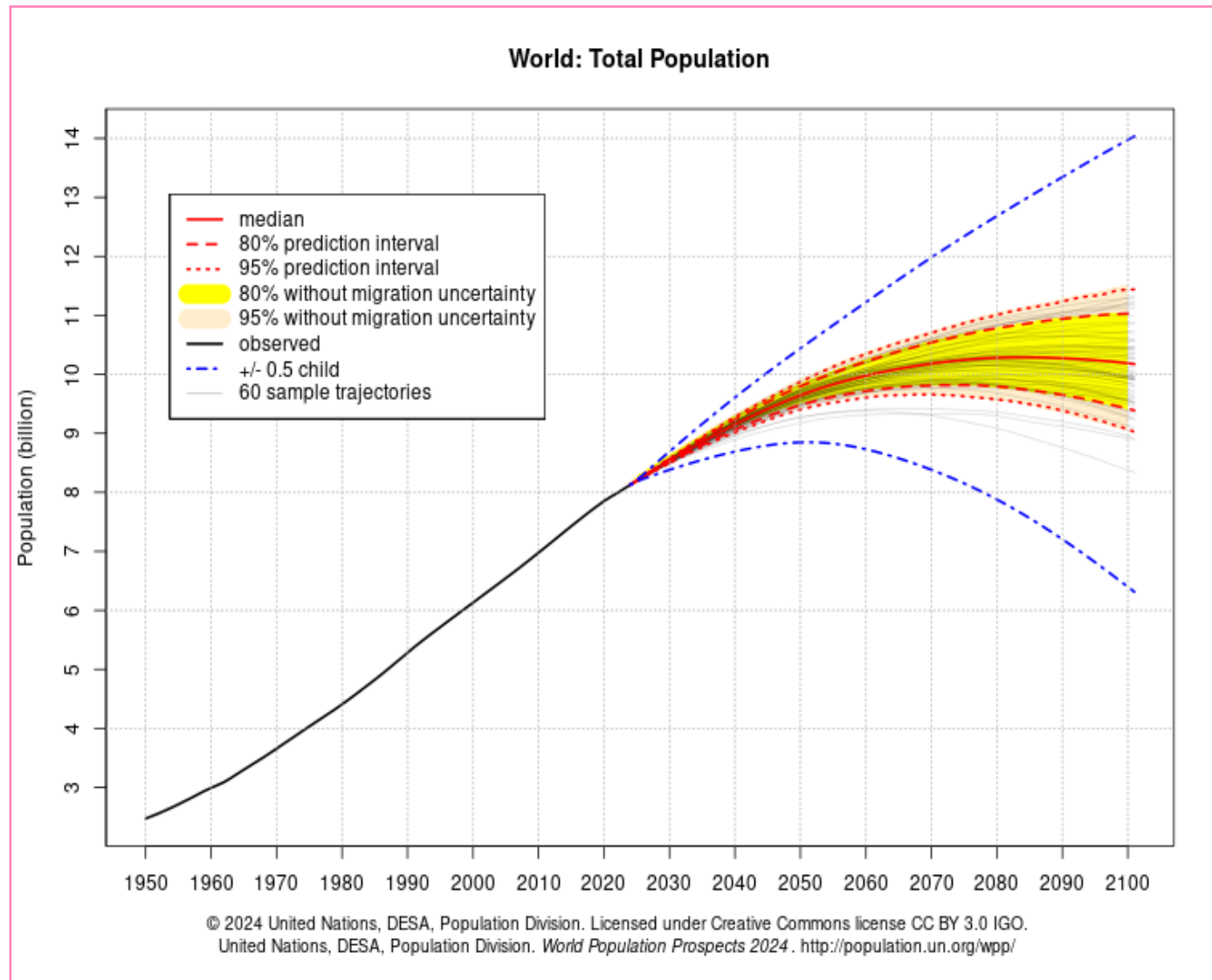
Сценарные и вероятностные прогнозы ООН

Первый (*средний*) вариант совпадает с медианой вероятностного прогноза.

Высокий вариант равен среднему варианту +0,5 рождений

Низкий вариант равен среднему варианту минус 0.5 рождений

Цель вариантов: показать, как изменения в рождаемости влияют на демографическую динамику



Почему прогнозы многовариантны?

- Мы определяем не «точечное значение» (вероятность попасть в «точное значение» равна нулю), а интервалы вероятных демографических изменений. В этом интервале ищем наиболее ожидаемую (вероятную) оценку будущего.
- Кроме того, в процессе построения реалистического прогноза могут решаться задачи аналитического характера, например, как сильная миграция влияет на численность молодежи и возрастной состав населения в целом.

Система демографических прогнозов отдела народонаселения ООН

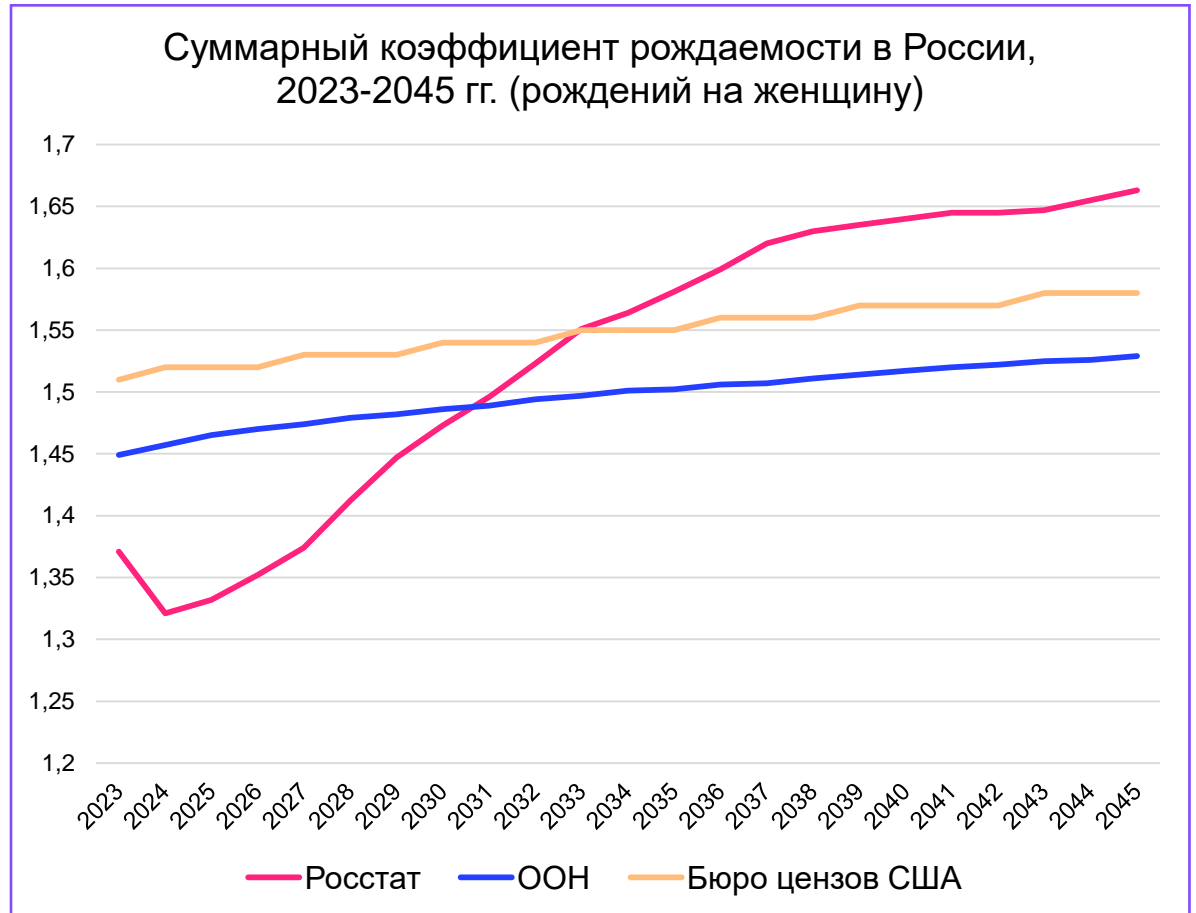
- Многовариантные прогнозы для всех стран мира, выполненные по единой методике.
- Прогнозы выполняются по принципу **ревизий (пересмотров)**: прежние прогнозы пересматриваются с учетом новых тенденций.
- Первый прогноз выполнен в 1951 г. Всего – 28 ревизий. Последняя ревизия - в 2024 г.
- В настоящее время прогнозы пересматриваются каждые два года и выполняются до 2100 г.
- Основа прогнозирования – созданная и поддерживаемая база данных о населении, в.т.ч. для стран, где информация неточная или отсутствует.
- Результаты в двух видах для каждой страны: (а) 13 вариантов сценарных прогнозов; (б) вероятностные прогнозы.
- Результаты, методология и данные представлены on-line: [//www.un.org/en/development/desa/population/](http://www.un.org/en/development/desa/population/)
- Долгосрочные прогнозы: до 2004 года выполнялись до 2100. В 2004 году выполнен прогнозы до 2300 года «**Foresight for Development**»

Варианты прогноза Отдела народонаселения ООН (2024)

		Сценарии	
название	Рождаемость	Смертность	Миграция
Средний	средний	средний	средний
Высокий (рождаемость)	[Средний] + 0,5 рождений	средний	средний
Низкий (рождаемость)	[Средний] - 0,5рождений	средний	средний
Постоянная рождаемость	Постоянная с 2024 г.	средний	средний
Рождаемость простого воспроизводства	Постоянная для простого воспроизводства	средний	средний
Нулевая ранняя рождаемость	[Рождаемость] = 0 до 18 лет	средний	средний
Ускоренное снижение рождаемости	Коэффициент рождаемости для женщин младше 20 лет снижается на 20% ежегодно до достижения уровня 10 рождений на 1000 женщин в группе 15–19 лет.	средний	средний
Ускоренное снижение рождаемости с восстановлением	Восстановление рождаемости на 50 процентов через 10 лет для соответствующих когорт	средний	средний
Постоянная смертность	Средний	Постоянная с 2024 г.	средний
Нет изменений	Постоянная с 2024 г.	Постоянная с 2024 г.	средний
Нулевая миграция	Средний	средний	Нулевая с 2024 г.
Простое воспроизводство с нулевой миграцией	Постоянная для простого воспроизводства	средний	Нулевая с 2024 г.
Моментум	Постоянная для простого воспроизводства	Постоянная с 2024 г.	Нулевая с 2024 г.

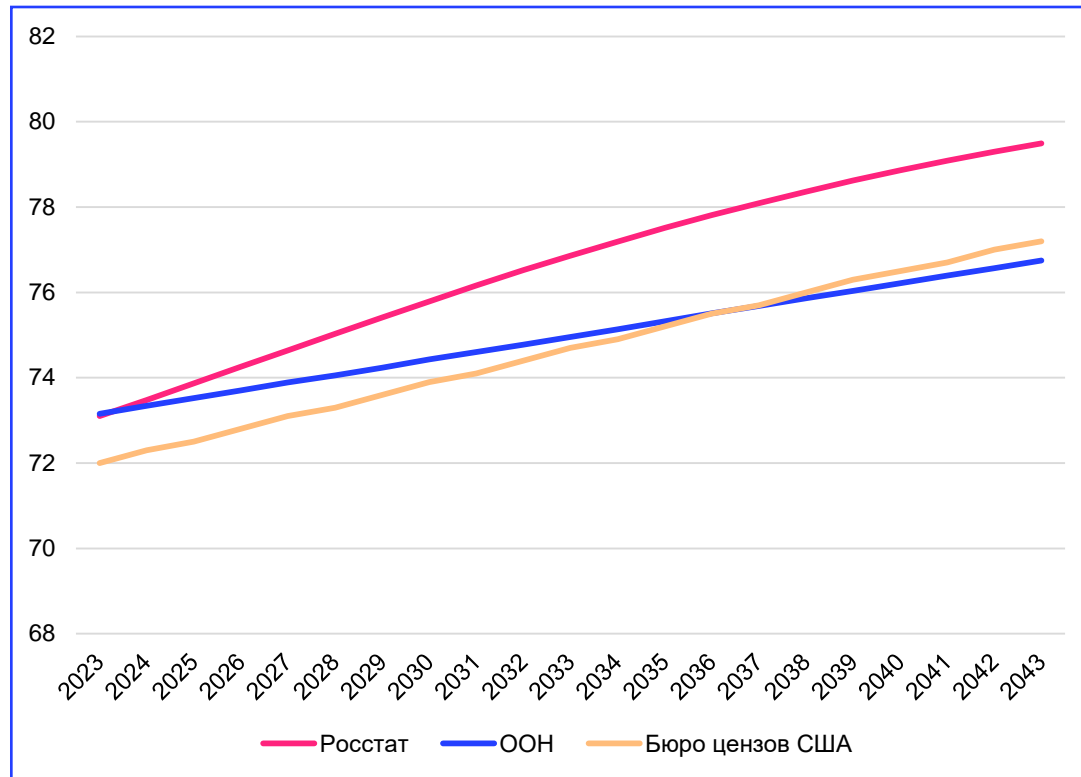
Почему нужны национальные прогнозы населения?

- Национальные прогнозы отличаются от прогнозов, выполненных для данной страны в международных организациях или в других государствах.
- Национальные прогнозы полнее учитывают специфику демографического развития и учета населения страны.
- Национальные прогнозы населения являются основой для построения прогнозов населения регионов страны.

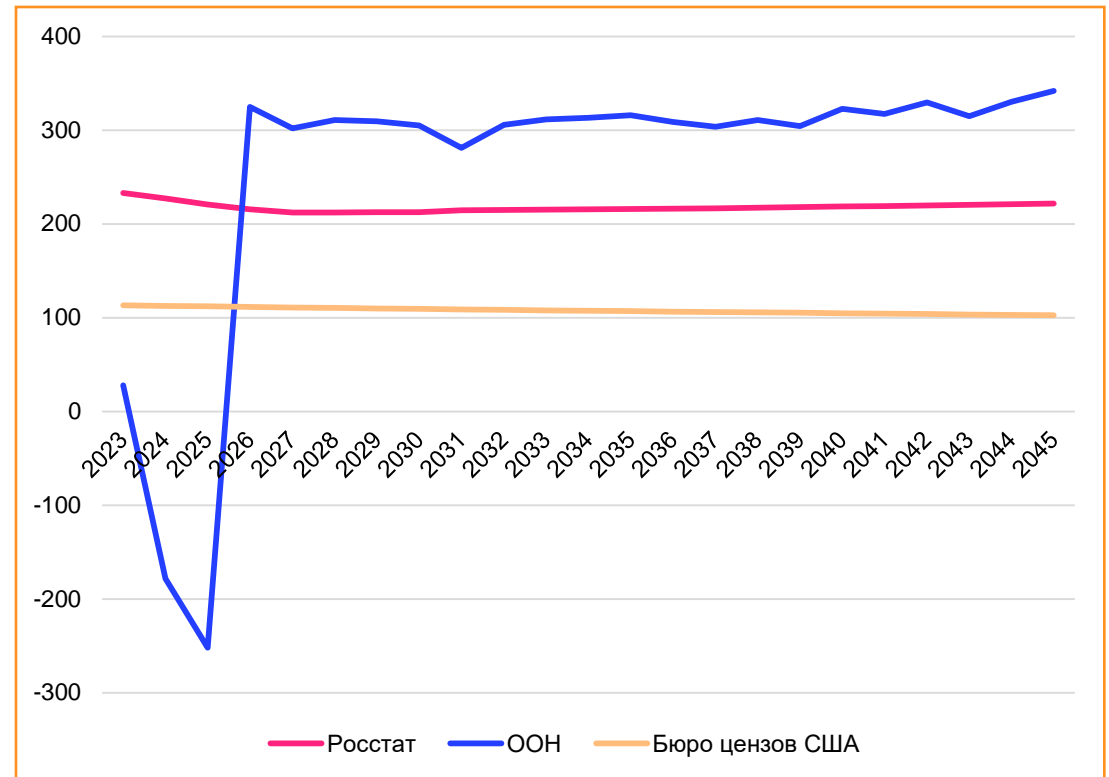


Прогнозы для России, построенные в Росстате, Отделе народонаселения ООН и Бюро цензов США (средние варианты)

Ожидаемая продолжительность жизни при рождении в 2023-2045 гг., лет (оба пола)




Миграционный прирост, 2023-2045 гг. (тыс. человек)





Точность демографических прогнозов

- Демографические прогнозы точны по сравнению с прогнозами экономического и социального развития. В основе точности – инерционность демографических изменений.
- В краткосрочной и среднесрочной перспективе демографические прогнозы точны (ошибка в оценке численности населения – несколько процентов, если нет каких-либо непредвиденных кризисных явлений).
- Наиболее точны прогнозы численности для поколений, которые жили на начальную дату прогноза.
- Большая часть ошибок в демографических прогнозах объясняется ошибками в сценариях рождаемости.
- Поэтому наименее точны для тех поколений, которые будут рождены в течение прогнозного периода.



Совокупная ошибка в прогнозах численности населения стран мира на 2000 год

Организация	Год	Ошибка, %
Всемирный Банк	1972	18
Отдел народонаселения ООН	1973	17
Всемирный банк	1983	13
ООН	1984	11
Бюро Цензов США	1987	12
Всемирный банк	1988	8
Всемирный банк	1990	7
Отдел народонаселения ООН	1994	4

Источник: Beyond six billions: forecasting the world's population.

J.Bongaarts and R.Bulatao (eds.). Washington, D.C.: National Academy Press 2000



МЕТОДЫ ПОСТРОЕНИЯ ДЕМОГРАФИЧЕСКИХ ПРОГНОЗОВ

Население без структуры

Когортно-компонентный метод

Методы прогнозирования демографических процессов

Основа любого прогноза

Анализ прошлых демографических тенденции и обуславливающих их факторов, а также изучение возможных изменений в факторах, влияющих на демографические процессы в будущем (введение новых мер семейной политики, перспективы экономического развития, новые технологии и пр.) - основа любого демографического прогноза.

(1) ПРОГНОЗЫ НАСЕЛЕНИЯ БЕЗ СТРУКТУРЫ ИЛИ ТРЕНДОВЫЕ МОДЕЛИ

Прогнозы изменения одного показателя, например, численности всего населения или отдельной социально-демографической группы, суммарного коэффициента рождаемости или ожидаемой продолжительности жизни

Определения

- Временной ряд – наблюдения одного объекта в разные моменты (периоды) времени
- Трендом называется постоянное долгосрочное изменение переменной во времени. Временные ряды колеблются вокруг своих трендов.
- В прошлом исследователи на основе ограниченных показателей рядов динамики строили «Законы роста населения».
- В настоящее время для выделения трендов и прогнозирования применяется современный статистический аппарат анализа временных рядов.

Прогнозы численности населения «без структуры»

Самый простой и исторически первый вид демографических прогнозов: прогноз общей численности населения или отдельных групп населения без разбивки по возрасту и полу.

Из истории:

- Г. Кинг (Англия, конец 17 в.): предсказывал, что численность населения Англии в 2000 г. в 8,3 млн. человек. Метод – арифметическая прогрессия;
- Т. Мальтус (Англия, конец 18 в.): алармистские прогнозы. Рост населения - в геометрической прогрессии (период удвоения – 25 лет);
- У. Уотсон (1810, США): прогноз численности населения США, экстраполяция коэффициентов роста;
- Д. Менделеев (Россия): численность населения Земли в 1965 г. – 10 млрд. человек вместо 3,3 млрд. Использовалась экстраполяция по формуле сложных процентов;
- С. Г. Струмилин (1920, СССР): прогноз до 1940 г. (177,1 млн. чел), гипотеза о постоянном коэффициенте прироста.
- Р. Пирл и Л. Рид: применение логистической кривой для прогноза населения США (1930, 1940) и Земного шара (1936) .
- Г. Гаузе (1930): логистическая модель СССР: 2000 г. - 193 млн. Европ. Часть СССР
- И. Шкловский (1965), С. Капица (1990-е гг.): Модель гиперболического роста населения мира

Основные типы трендов:

- ☐ Линейный
- ☐ Параболический
- ☐ Экспоненциальный
- ☐ Логистический
- ☐ Гиперболический

Пример линейной функции

Условие

$P(t)$ – население в момент ;

$P(0)$ – начальное население;

t – время (известно);

d - прирост населения.

$$P(t) = P(0) + (t-1) \cdot d$$

Гипотеза r – коэффициент прироста населения

Коэффициент изменяется линейно:

$$r = a + b \cdot t$$

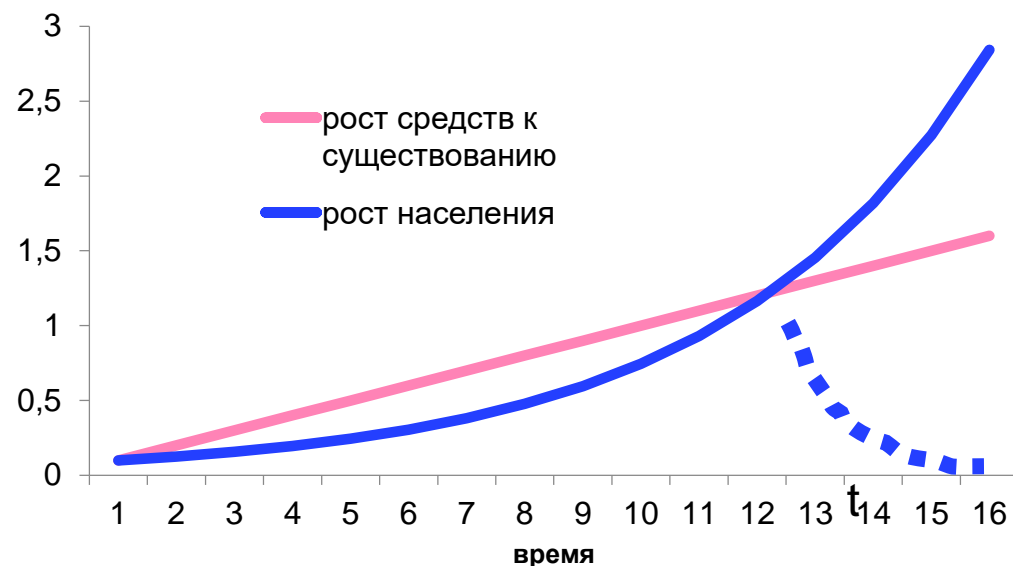
Модели демографического роста в случае постоянных темпов прироста r

Экспоненциальная модель (рост не зависит от плотности населения)

$$P(t) = P(0) * e^{rt}$$

$P(0)$ – начальное население;

$P(t)$ – население в момент t ;

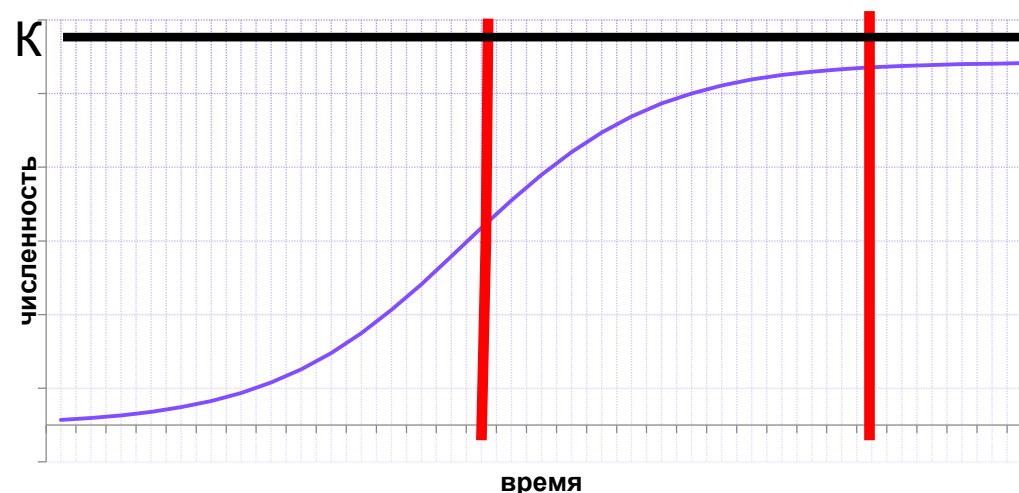


Логистическая модель (рост зависит от плотности населения)

$$P(t) = \frac{K}{1 + e^{\alpha - rt}}$$

r - коэффициент прироста населения

K (емкость территории), α , – параметры уравнения



Параболическая модель

$$P(t) = a_0 + a_1t + a_2t^2 + \dots + a_nt^n$$

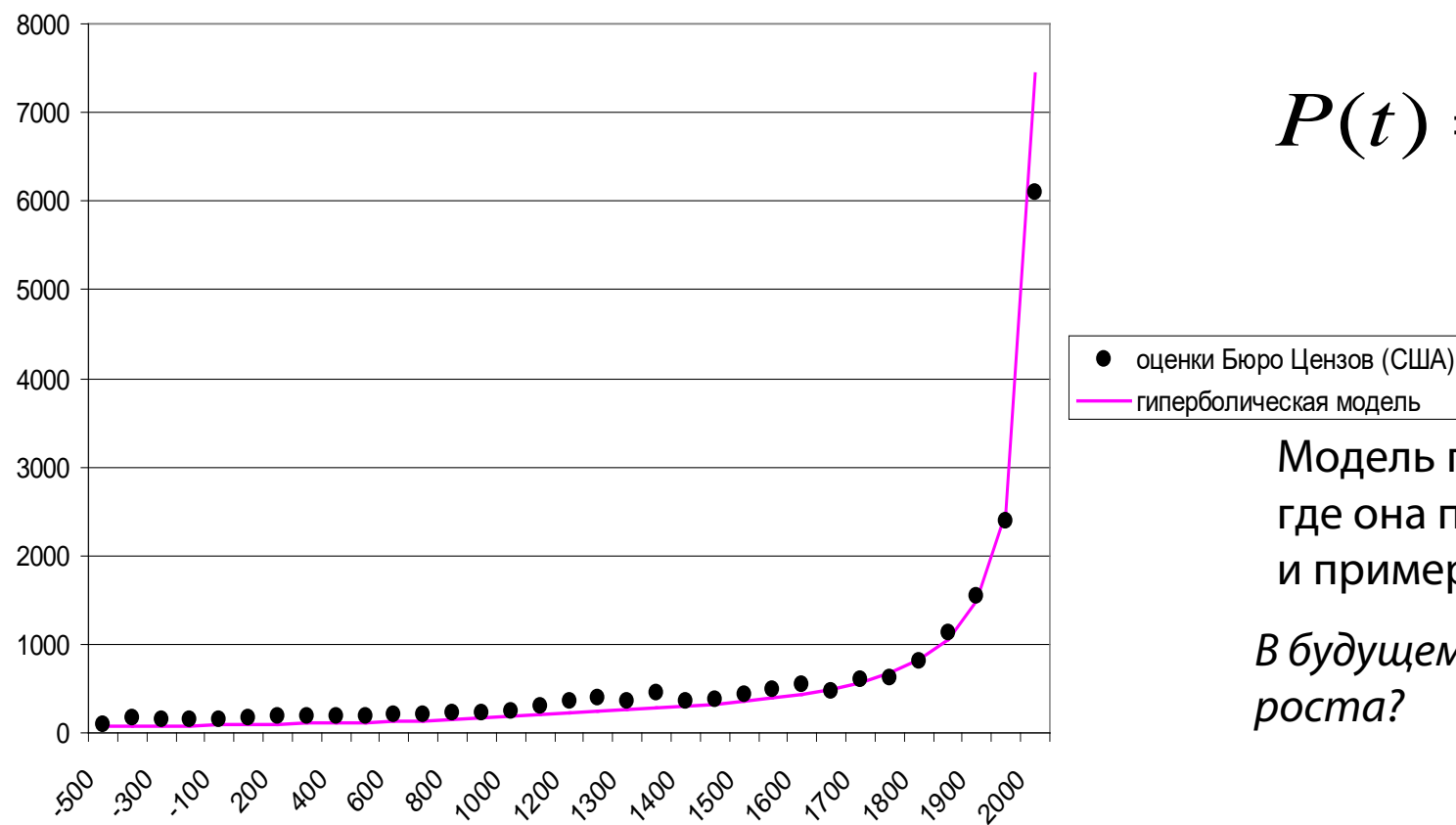
Где $P(t)$ – численность населения в момент t ;

t – независимая переменная время;

a_i параметры.

На практике используют, как правило, параболы второго ($n=2$) и третьего ($n=3$) порядков

Гиперболическая модель роста мирового населения С.П.Капицы



$$P(t) = 186 \cdot 10^9 / (2025 - t)$$

Модель принципиально ограничена областью, где она применима (начало истории человека и примерно до 2025 г.).

В будущем сменяется логистической моделью роста?

Капица С.П. Теория роста населения Земли. М., 1997.

Меняющаяся скорость демографического роста

Пусть коэффициенты прироста населения меняются на протяжении периода $(0, T)$ таким образом, что на каждом временном интервале Δt ($\sum \Delta t = T$) наблюдаются различные, но постоянные на этих интервалах темпы прироста населения r_0, r_1, r_2, \dots . Тогда численность населения P в момент T равна:

$$\begin{aligned} P(T) &= P(0) \cdot e^{r_0 \cdot \Delta t} \cdot e^{r_1 \cdot \Delta t} \cdot e^{r_2 \cdot \Delta t} \dots = \\ &= P(0) \cdot e^{r_0 \cdot \Delta t + r_1 \cdot \Delta t + r_2 \cdot \Delta t + \dots} = P(0) \cdot e^{\int_0^T f(r) dt} \end{aligned}$$

Развитие трендовых моделей

- Главный недостаток трендовых моделей при прогнозировании развития сложных объектов, состоящих из разных элементов, таких как население, является то, что они не учитывают их структуру.
- В настоящее время трендовые модели в виде моделей временных рядов используются для прогнозирования показателей интенсивности демографических процессов (см. ARIMA, Lee-Carter и др.).
- Пандемия COVID -19 повысила интерес к временным рядам с сезонной и циклической компонентами.

Важно!

Если будущее отличается от прошлого, то взаимосвязи, оценённые на основе исторических данных, не очень полезны при прогнозировании будущего.

Прогнозы - это не только тренды. Тренды не могут оставаться постоянными. Следует учитывать возможные изменения, по крайней мер в ближайшей перспективе, в образовании, охране здоровья, уровне жизни и пр.



(2) КОГОРТНО-КОМПОНЕНТНЫЙ МЕТОД

Прогноз численности населения и его возрастно-половых групп, чисел родившихся, умерших и мигрантов

Впервые применен на рубеже 1920-1930-х гг.

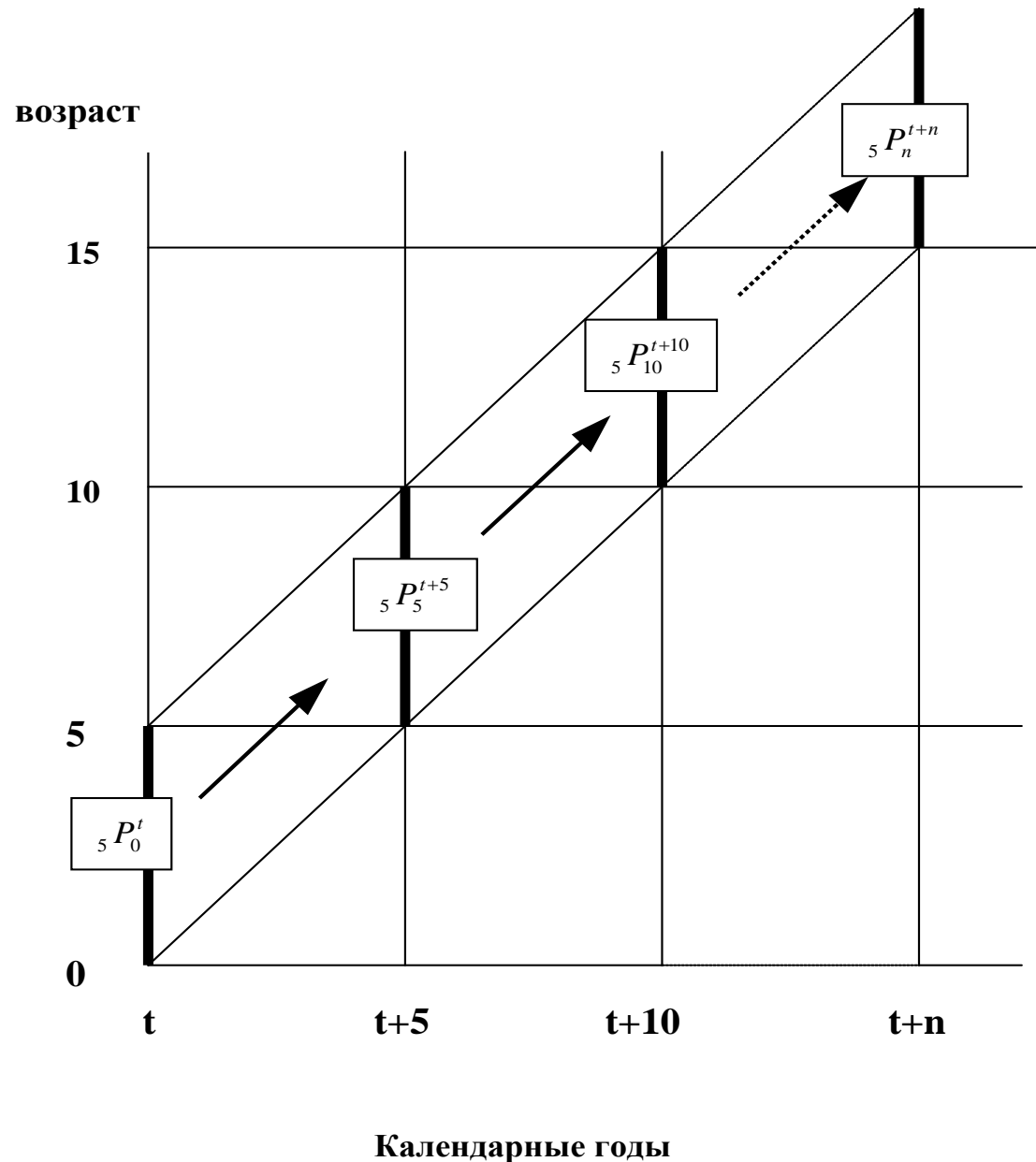
- На 19-й сессии международного статистического института (Токио) четыре доклада по перспективным исчислениям населения: К.Джини (Италия), А.Иенсен (Дания), М.Птуха (СССР), У.Томсон (США).
- Демографический институт Академии наук УССР, М.Птуха Население Украины до 1960 г. Бюллетень международного статистического института, Гаага, 1931 г. (на фр. языке)
- Демографический институт АН СССР в г. Ленинграде, В. В. Паевский. О методах расчета будущей численности и возрастного состава населения. "Доклады Академии наук СССР". 1931, № 2, стр. 45-49.
- В США: W.S. Thompson and P.K. Whelpton 1933. Population Trends in the United States.).
- Лига наций: 1944 год (прогноз – предшественник системы прогнозов Отдела народонаселения ООН)

Когортно-компонентный метод

- Это реализация дискретной модели динамики населения. Характеристики населения вычисляются только для определенных (опорных) моментов времени.
- Шаг прогнозирования равен длине возрастного интервала прогнозируемого населения
- Оценки численности возрастно-половых групп в иные моменты времени получаются путем **интерполирования** результатов прогноза между ближайшими точками, для которых выполнялись прогнозные оценки.

Почему метод называют когортным?

Потому что
прогнозируется
численность когорт.
Рассчитывается
численность каждой
когорты по мере ее
перехода из одной
возрастной группы в
другую



Почему метод называют компонентным?

Потому что оценки численности когорт на основе и всего населения лежат компонент демографического роста, которые в итоге объединяются в уравнение демографического баланса следующим образом:

$$P_{t+n} = P_t + V_{t,t+n} - D_{t,t+n} + \Delta M_{t,t+n}$$

где

$V_{t,t+n}$ – число родившихся за период $(t,t+n)$

$D_{t,t+n}$ – число умерших за период $(t,t+n)$

$\Delta M_{t,t+n}$ – нетто-миграция за период $(t,t+n)$

n - шаг прогнозирования

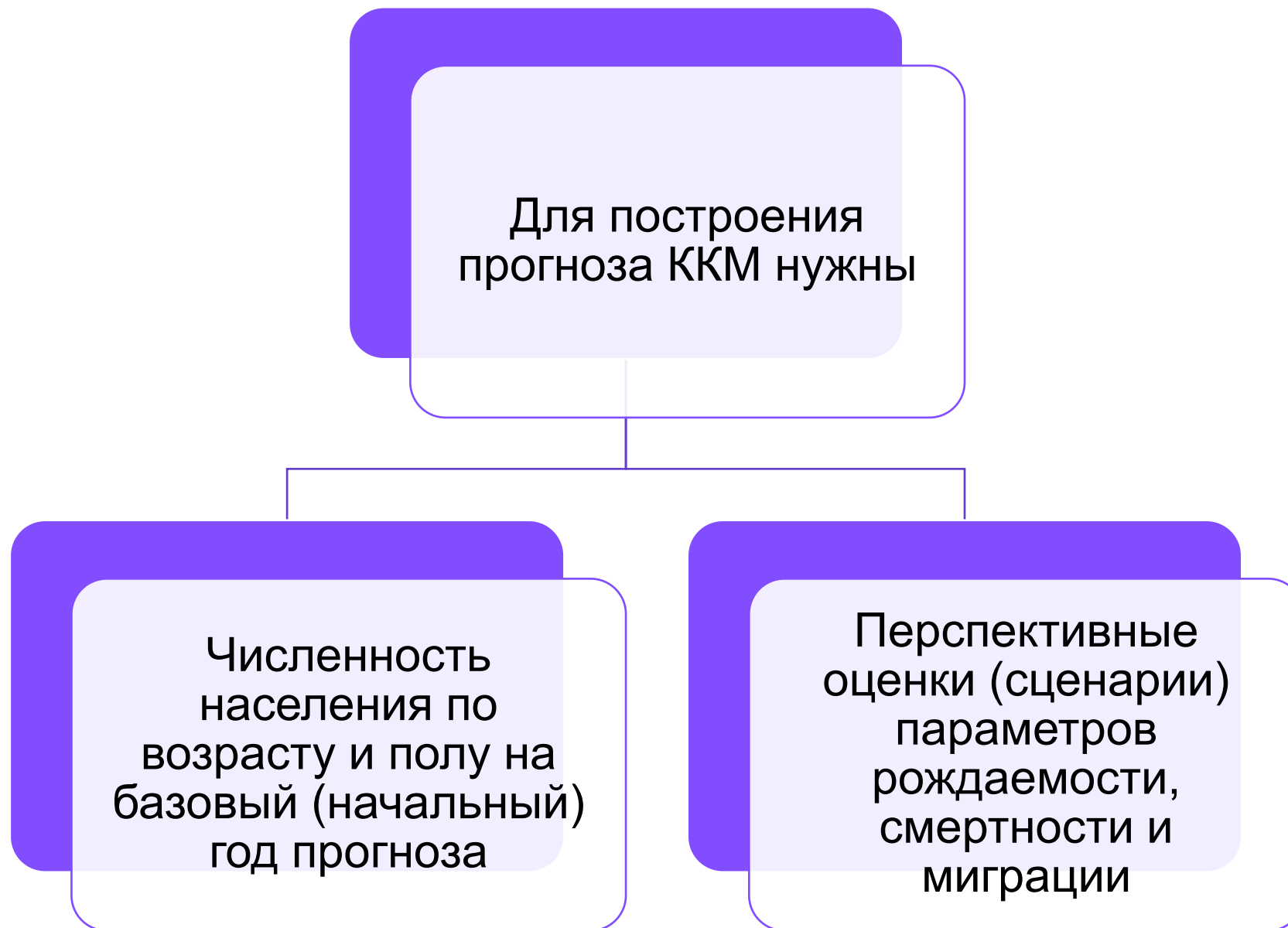
Почему метод называется
компонентным?

**Численность группы в возрасте $x+n$ старше 1
года:**

$$P_{t+n}(x+n) = P_t(x) - D_{t,t+n}(x) + \Delta M_{t,t+n}(x)$$

**Численность начальной возрастной группы (до
1 года):**

$$P_{t+n}(0) = B_{t,t+n} - D_{t,t+n}(0) + \Delta M_{t,t+n}(0)$$



Каждый шаг прогнозирования можно разбить на 4 стадии

1. Передвижка когорты с учетом смертности в следующую возрастную группу.
2. Вычисление числа родившихся за шаг прогнозирования.
3. Вычисление численности первой возрастной группы (дошедшие до следующего момента времени из числа родившихся).
4. Включение иммигрантов и эмигрантов (или миграционного прироста) с учетом родившихся и умерших среди них.



Разработка сценариев (гипотезы) изменений демографических процессов – главный этап построения прогнозов ККМ:

Речь идет о сценариях (прогнозах) рождаемости, смертности и миграции.

Для реализации когортно-компонентного метода необходимы показатели уровня и возрастные характеристики рождаемости, смертности и миграции.

Этапы реализации демографического прогноза

1. Постановка задачи, выбор метода прогнозирования и спецификация параметров прогностической модели.
2. Сбор, коррекция и анализ данных.
3. Разработка прогнозных сценариев изменений рождаемости, смертности и миграции.
4. Ввод данных и реализация прогноза.
5. Оценка и анализ результатов. При необходимости переход к этапу 3 и уточнение сценариев.
6. Представление результатов.

Анализ демографических тенденций – основа разработки прогнозных сценариев

Разработке прогнозных сценариев должен предшествовать комплексный углубленный анализ демографического развития, который включает в себя решение следующих задач:

- анализ долговременных тенденций демографических процессов, предшествующих году начала реализации прогноза;
- анализ текущей ситуации в области рождаемости, смертности и миграции;
- выявление причин, обуславливающих изменения в демографической ситуации;
- оценка возможных направлений в изменении рождаемости, смертности и продолжительности жизни, миграционных процессов.

Когортно-компонентный подход для вероятностных прогнозов

Сценарии демографических процессов разрабатываются в формате случайных функций. Когортно-компонентный метод в данном случае представляет собой имитационную модель. Каждая имитация есть независимый прогноз для комбинации демографических переменных, возникающих в случайном порядке. В результате получаем не единственную траекторию, а интервал траекторий.

Задачи по подготовке данных:

- Для разработки сценариев необходимо собрать надежные данные за максимально длительный период времени, чтобы выявить тенденции в изменении демографических процессов
- Базовая возрастная структура должна быть тщательно проверена, все деформации, обусловленные ошибками в сборе данных, должны быть устранены (сглажены).
- Базовые данные для построения таблиц смертности должны быть выверены, все искажения и неровности в кривых смертности, обусловленные ошибками в сборе данных, должны быть устранены (сглажены).
- Базовые данные для оценки возрастных коэффициентов рождаемости должны быть тщательно проверены и, при необходимости, скорректированы.
- Особое внимание уделить оценкам миграции (полнота учета).
- Особое внимание (об этом часто забывают) следует уделить оценкам соотношения численности полов при рождении (особенно в тех странах, где оно сильно отклоняется от среднемировых показателей).

Переход от национального прогноза к региональным, и наоборот - от региональных к национальному

Простой метод

Требуют минимум информации.

Результаты национального прогноза распределяется между регионами по определенным правилам. Например, на основе анализа изменений доли каждого из регионов в общей численности населения страны и отдельных возрастных групп.

Когортно-компонентный метод

выполняются национальный прогноз и прогнозы для регионов с учетом межрегиональной миграции (миграционный прирост за счет внутренней миграции). Полученные результаты для страны в целом и суммы регионов балансируются по определенным правилам.

Сложная версия: построение мульти статусной (много региональной) модели, в которой представлены миграционные потоки между регионами.

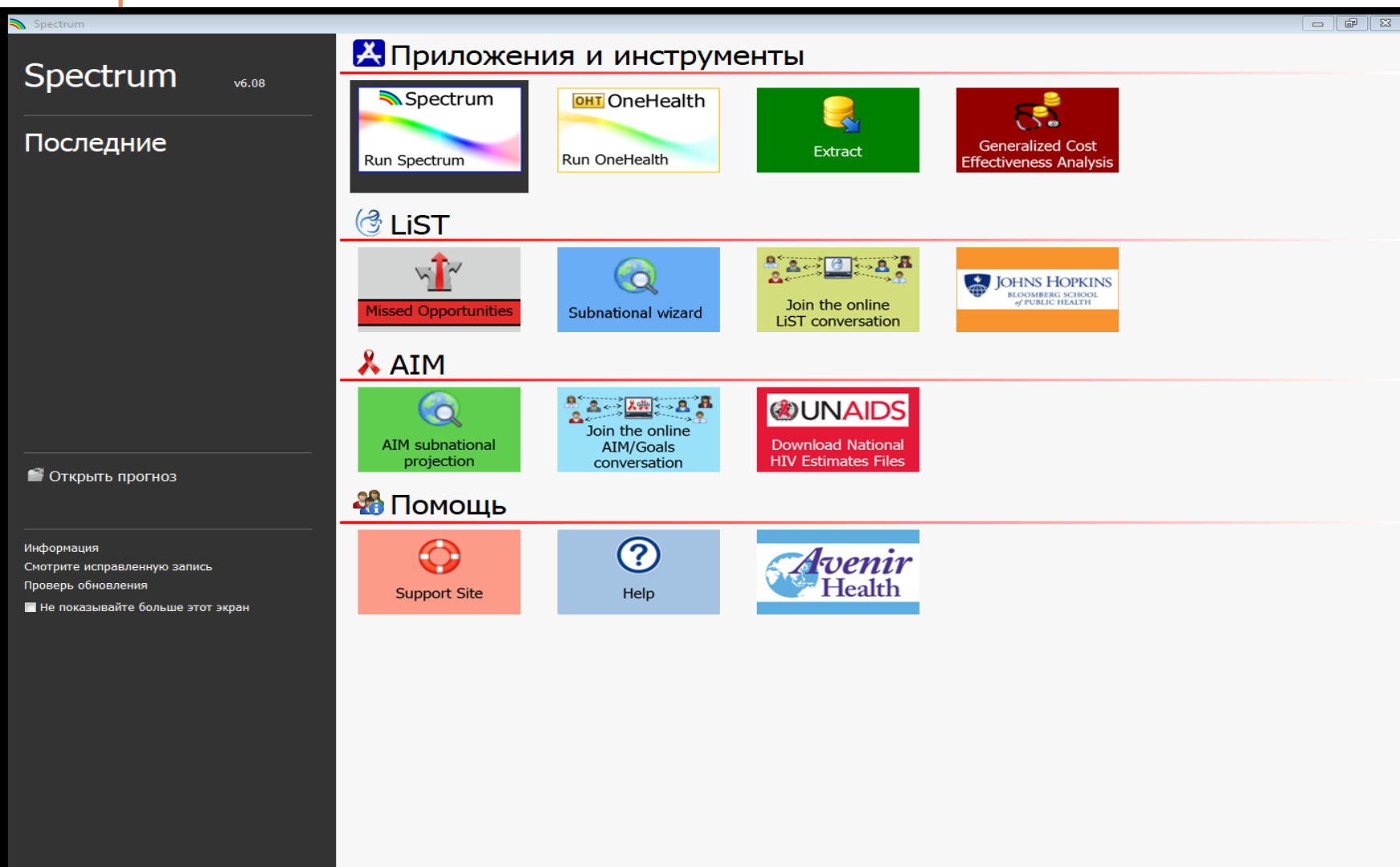
Мультистатусные в т.ч. многорегиональные прогностические модели

- В основе многорегиональных моделей – вероятности перехода из одного региона в другой с учетом возраста и пола мигрантов.
- Модель является попыткой наиболее точного описания демографического развития территориальной системы, состоящей из N-административно-территориальных единиц.
- Среди показателей модели, например, продолжительность жизни с учетом продолжительности проживания в том или ином регионе.
- Недостатки: (1) модель громоздка. Требует большого количества данных. (2) модель чувствительна к качеству данных. В случае «плохих» данных по миграции модель становится неустойчивой. (3) Относительно хорошо работает, если регионов или страт немного. Например, крупные провинции (Канада) или образовательные группы (в случае мультистатусных моделей).



(3) ДОСТУПНОЕ (БЕСПЛАТНОЕ) ПРОФЕССИОНАЛЬНОЕ ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

Spectrum



Spectrum – пакет интегрированных программ, предназначенных для оценки последствий демографических изменений. Разрабатывается AVENIR Health при поддержке ряда международных и национальных организаций. Программа состоит из 2-х модулей с 18 подмодулями. Особый акцент сделан на решении задач,

MORTPAK FOR WINDOWS, Version 4.4



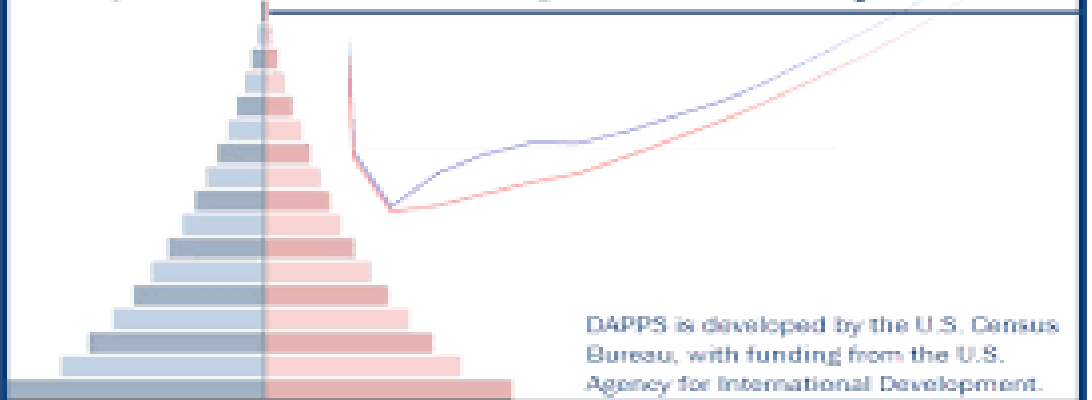
Популярный пакет для проведения демографического анализа и прогнозирования, включающий 17 приложений, в том числе для построения демографического прогноза, построения таблиц смертности, подбора модельных таблиц смертности. Непрямых методов оценки уровней смертности и рождаемости, оценки параметров модели стабильного населения и др. Первая версия разработана в Отделе народонаселения ООН в 1988 г.

RUP (Rural and Urban Projections) и DAPPS (Demographic Analysis and Population Projection System)

- DAPPS фактически является оболочкой RUP.
- Совместима с рядом программ по демографии, в том числе MORTPAK, и с Excel
- Профессиональная программа, позволяющая реализовывать как реалистические, так и аналитические прогнозы
- Хорошая визуализация и вывод данных.
- Возможность выполнять Многорегиональные прогнозы одновременно для 50 территорий.
- Балансирует прогностические сценарии для много региональных прогнозов.

Программы разработана Бюро Цензов США, и предназначены для выполнения межпереписных и постпереписных оценок населения, оценки качества данных и выполнения демографических прогнозов.


Demographic Analysis and Population Projection System





РАЗРАБОТКА ПРОГНОСТИЧЕСКИХ СЦЕНАРИЕВ

Главная и самая сложная фаза демографического прогноза



Разработка сценариев (гипотез) изменений в рождаемости, смертности и миграции— главный этап построения демографического прогноза когортно-компонентным методом

Для каждого демографического процесса прогнозируются:

- параметры изменения масштаба процесса (уровень рождаемости, смертности и миграции),
- параметры изменение возрастной модели процесса (рождаемости, смертности, миграции).



Методы прогнозирования демографических процессов

- Статистический анализ временных рядов (трендовые модели)
- Демографические модели
- Метод аналогий (или «таргетирование»)
- Структурные (факторные) модели демографических процессов
- Метод экспертов
- Байесовские модели (подход)
- Опыт ошибок прошлых прогнозов
- Учет национальных прогнозов (для регионов)
- Нормативы (планы) развития страны и регионов.

Выбор метода определяется (а) перечнем прогнозируемых показателей, (б) качеством имеющихся данных, (в) характеристиками самого выбранного метода.

Методы оценки вероятностных сценариев изменений демографических процессов

1. Метод экспертов
2. Статистический анализ временных рядов
3. Статистический анализ ошибок прошлых прогнозов

Главное условие для успешной разработки прогностических сценариев

- **Наличие команды из демографов профессионалов.**

- Кроме того:

Наличие качественных данных

И, что важно в современных условиях: **наличие программного обеспечения** с реализацией современных методов прогнозирования



(А) ВРЕМЕННЫЕ РЯДЫ

Прогнозирование начинается с выделения трендов.

Прогнозы населения, в которых предполагается постоянство показателей рождаемости, смертности и миграции, также имеют аналитическую ценность.

Модель Ли-Картера (Lee-Carter Model)

Идея модели заключается в разложении возрастных коэффициентов смертности $M(x,t)$ на две компоненты: изменяющегося со временем показателя $k(t)$ и набора возрастных констант $b(x)$:

$$\ln[M(x, t)] = a_x + K(t) \cdot b_x + \varepsilon(x, t)$$

где $t=1,2\dots T$, $x=1,2\dots n$

Параметры модели имеют следующую интерпретацию:

- Постоянная $a(x)$ отражает устойчивую во времени возрастную составляющую смертности;
- Постоянная $b(x)$ описывает колебания смертности в возрасте x как реакцию на изменения в общем уровне смертности;
- $K(t)$ – переменная, отражающая тренд общего уровня смертности.

Анализ по ряду стран с надежной статистикой показал, что $K(t)$ снижалась в 20 веке практически линейно, как и ожидаемая продолжительность жизни.

Метод используется моделирования и прогнозирования возрастных коэффициентов смертности и ожидаемой продолжительности жизни. Получив оценки для $k(t)$, и, затем, экстраполируя их, получаем оценки возрастных коэффициентов смертности $m(x,t)$.

Реализация модели

- Параметры модели рассчитываются с помощью метода SVD (Singular Value Decomposition, Сингулярное разложение), снижающего размер исходной матрицы.
- Линейный тренд $K(t)$ положен в основу прогностической модели (Random Walk with Drift)

$$K(t + 1) = \alpha + K(t) + \varepsilon(t),$$

Где α – среднее ежегодное изменение K (линейный дрейф)

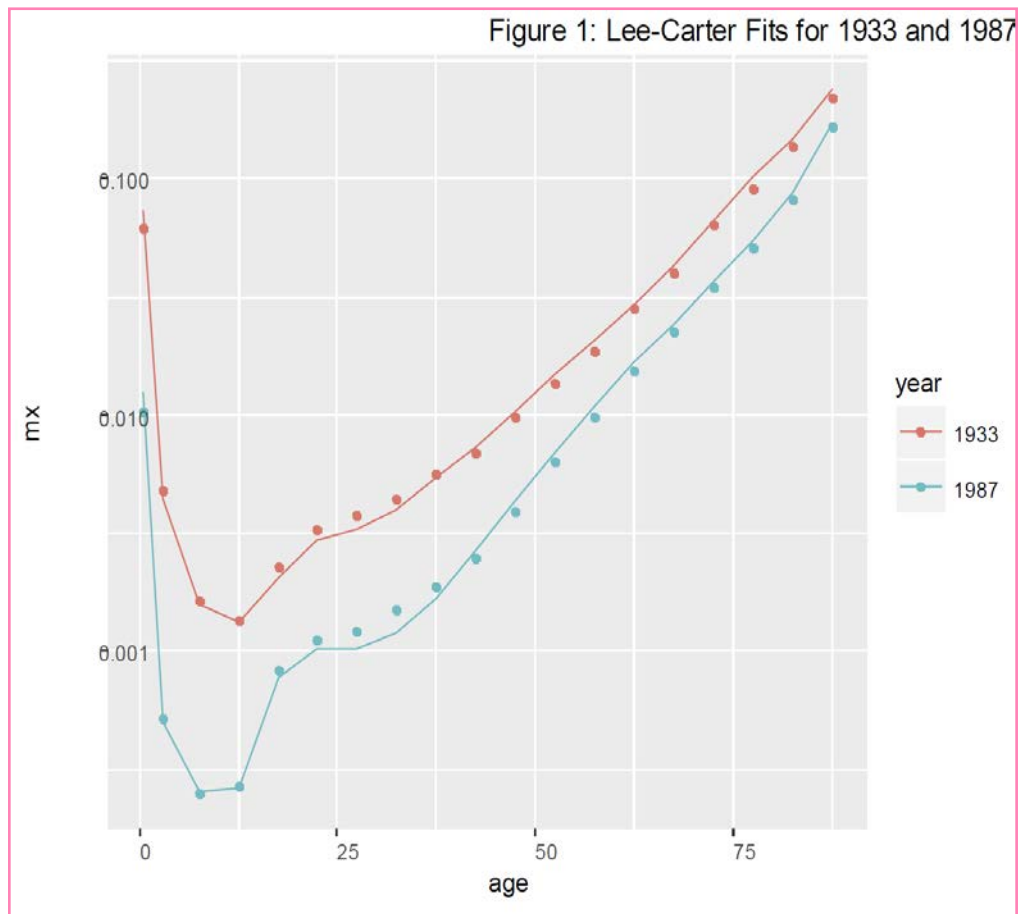
$\varepsilon(t)$ – случайная ошибка

Это модель случайного блуждания $\varepsilon(t)$ с дрейфом α .

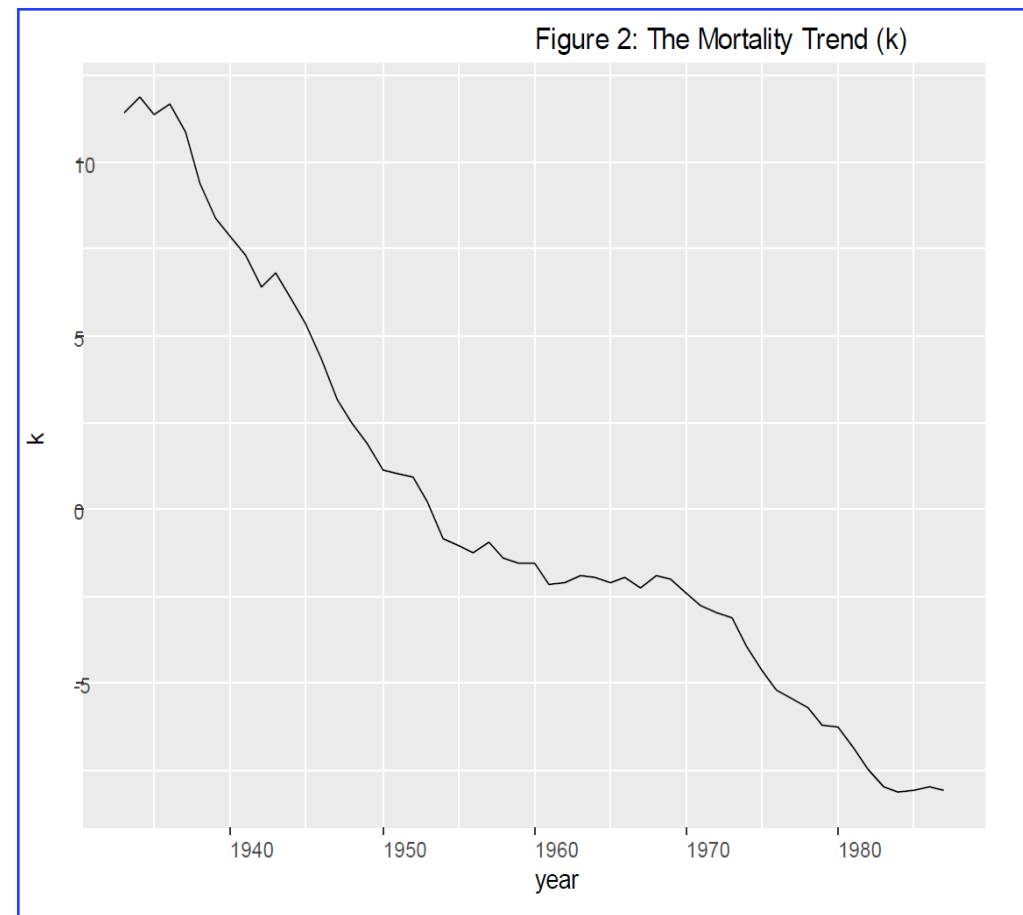
Все расчетные процедуры реализованы в STATA и R.

Изменение параметров модели ЛИ-Картера (на примере США)

Возрастная модель



Динамика $K(t)$. США 1933-1987.

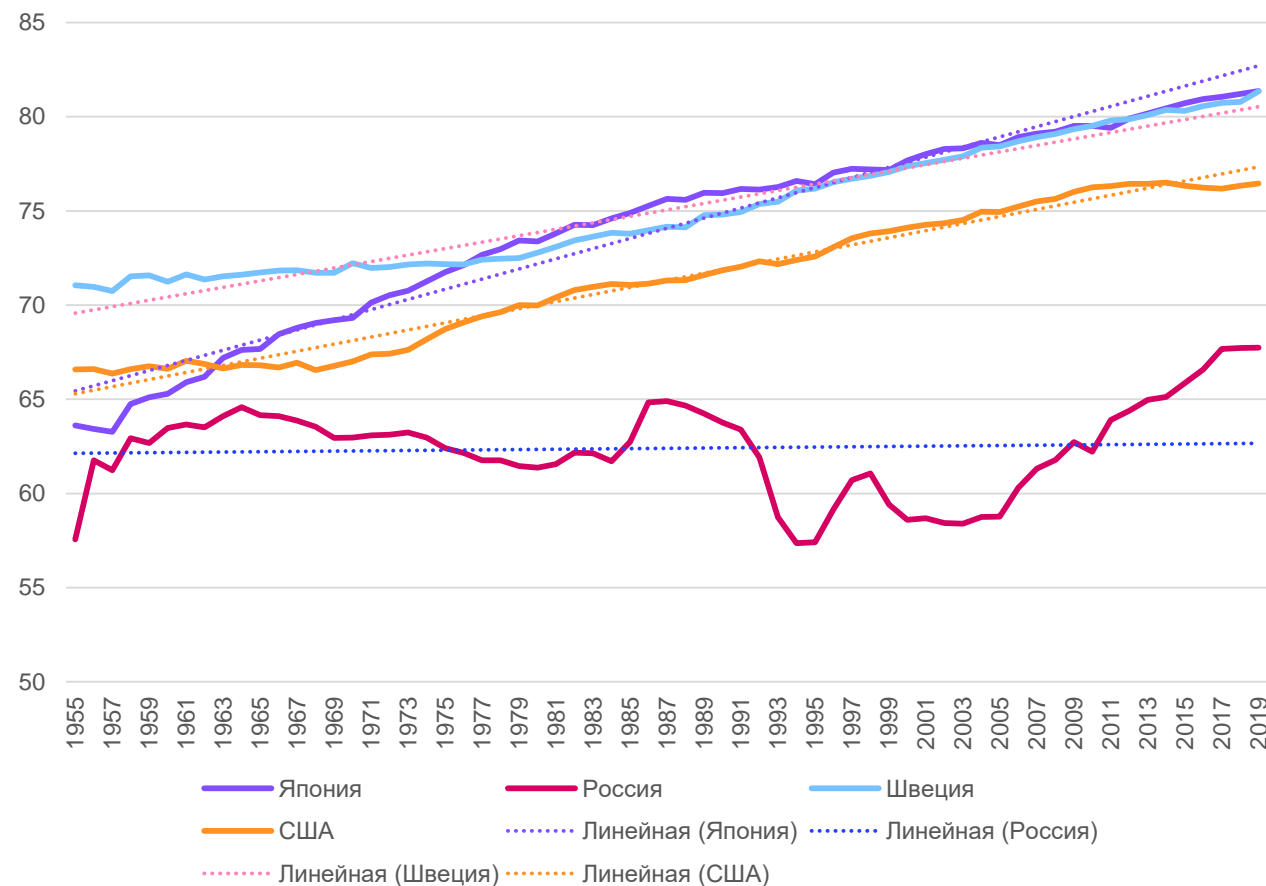


Ли-Картер - не панацея

Критика гипотезы линейности $K(t)$

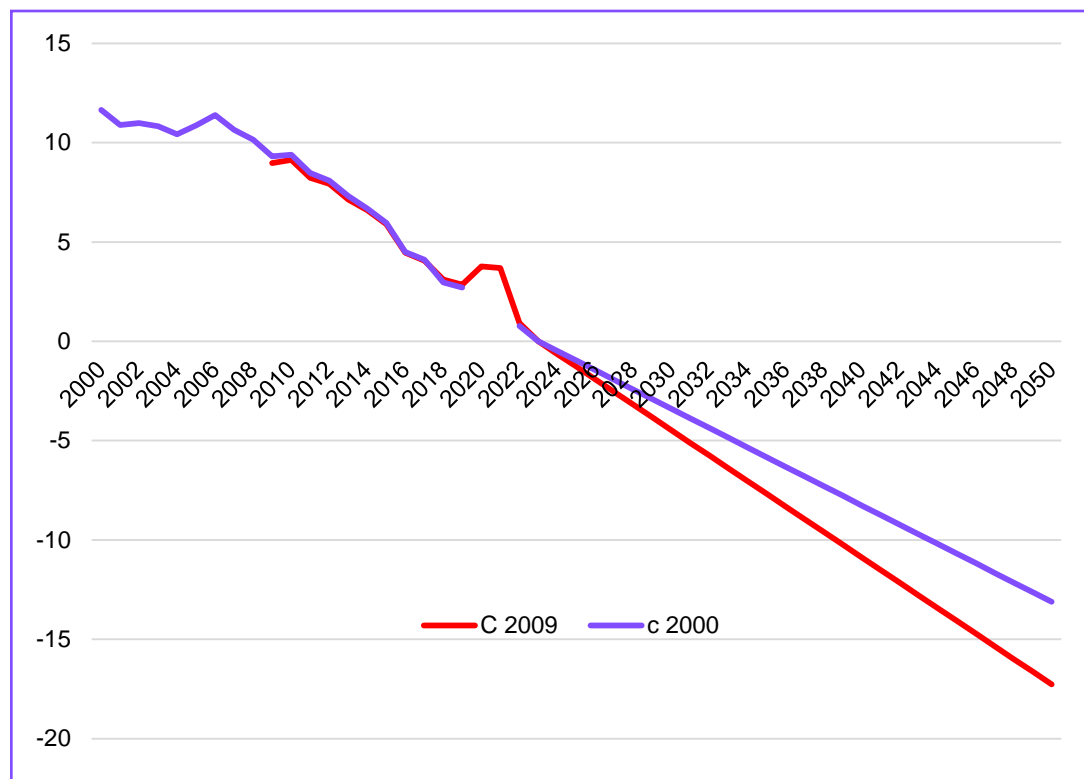
- Характерна для стран Запада до начала 21 века.
- Во многих странах линейность наблюдалась на коротких исторических периодах
- Возрастные изменения в настоящее время неравномерны по возрастам, поэтому параметр b_x нельзя считать постоянной величиной
- Результаты чувствительны к выбору интервала для оценки параметров в условиях нестабильного роста ОПЖ.

Ожидаемая продолжительность жизни (ОПЖ) при рождении, мужчины Японии, Швеции России и США.

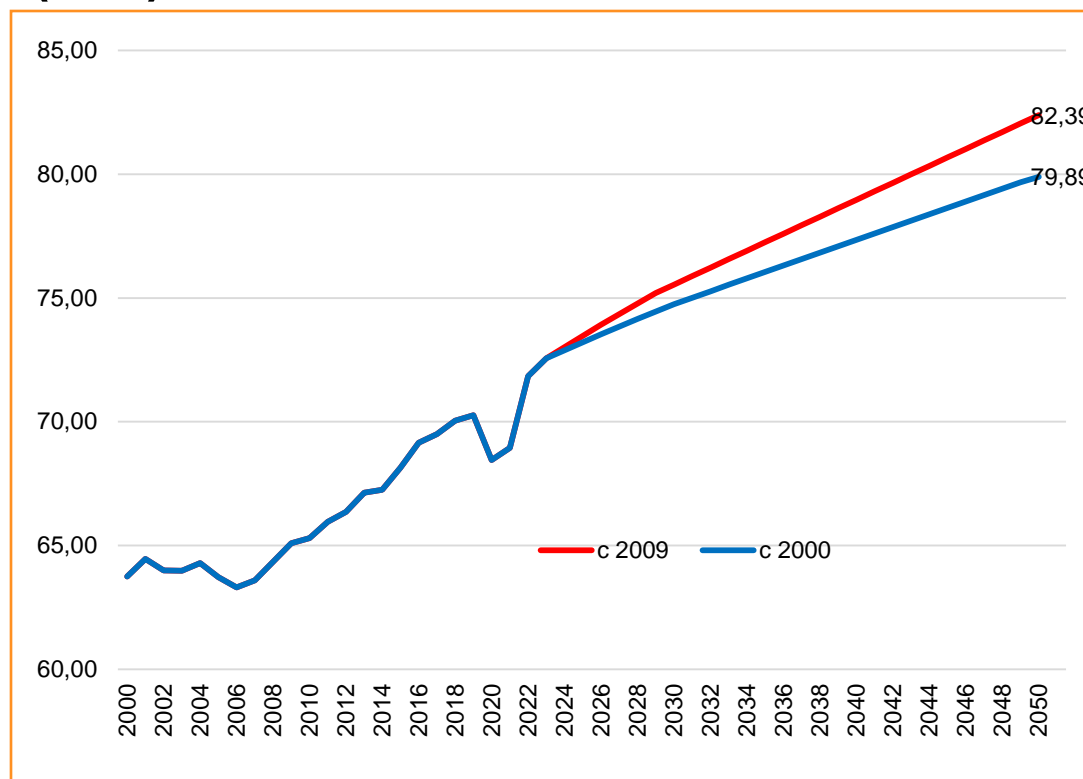


Результаты применения модели Ли-Картера к данным Кыргызстана: оценки с 2000 и с 2009 года.

Параметр $K(t)$ для мужского населения



Ожидаемая продолжительность жизни при рождении, мужчины (лет)



Примечание: ковидные годы не учитывались при моделировании

Авторегрессия AR. Модель ARIMA

- Модель авторегрессии AR(p): значения переменной Y_t в будущем зависят от ее значений в прошлом (запаздывание).
$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \dots + \beta_p Y_{t-p} + u_t.$$
 P – порядок авторегрессии
- Если ряд нестационарный, то для его превращения в стационарный берётся разница $\Delta y_t = y_t - y_{t-1}$. Получаем интегрированную авторегрессионную модель. Параметр d – количество разностей, необходимых для получения стационарного ряда.
- Интегрированная модель авторегрессии с ошибками в форме скользящего среднего (ARIMA(p,d,q)-модель):

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \dots + \beta_p Y_{t-p} + \theta_{t-1} u_{t-1} + \dots + \theta_{t-q} u_{t-q} + u_t$$

Параметр q – количество лагов ошибок. Моделирует зависимость текущего значения от прошлых ошибок прогноза

Применение модели ARIMA в демографии

Модель ARIMA популярна. Этому способствует ее реализация в многочисленных статистических пакетах и приложениях (STATA, R, GRETL и др.).

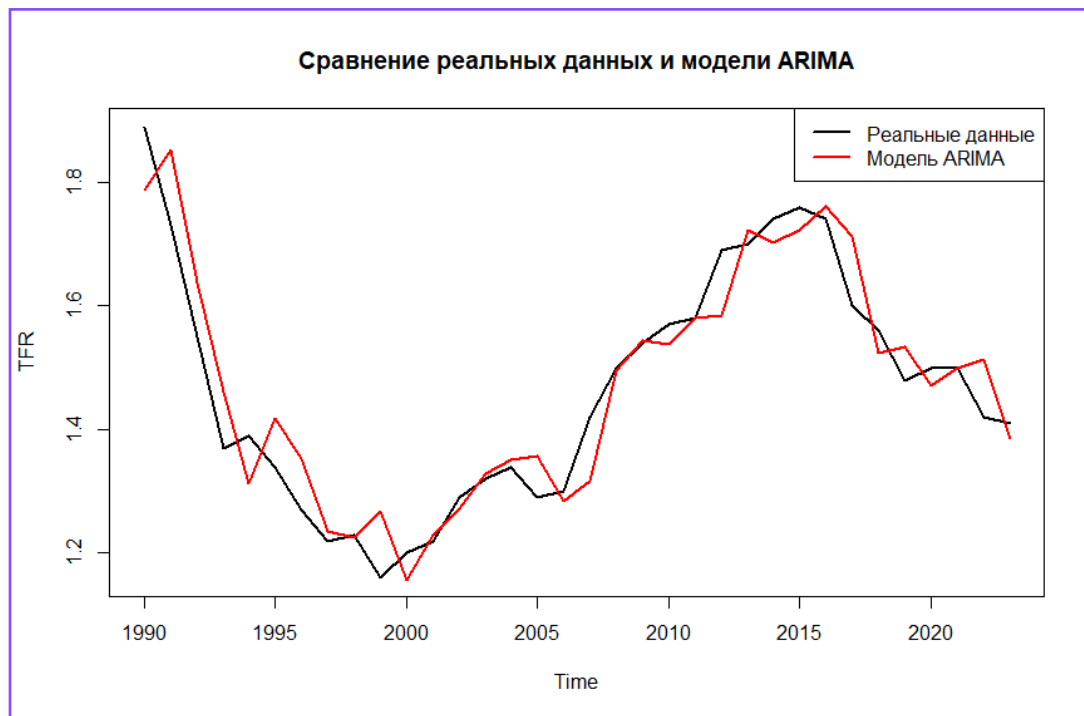
- Модель Ли-Картера – частный случай ARIMA (0,1,0):

$$K(t+1) = \alpha + K(t) + \varepsilon(t) \text{ или } \alpha = K(t+1) - K(t) + \varepsilon(t)$$

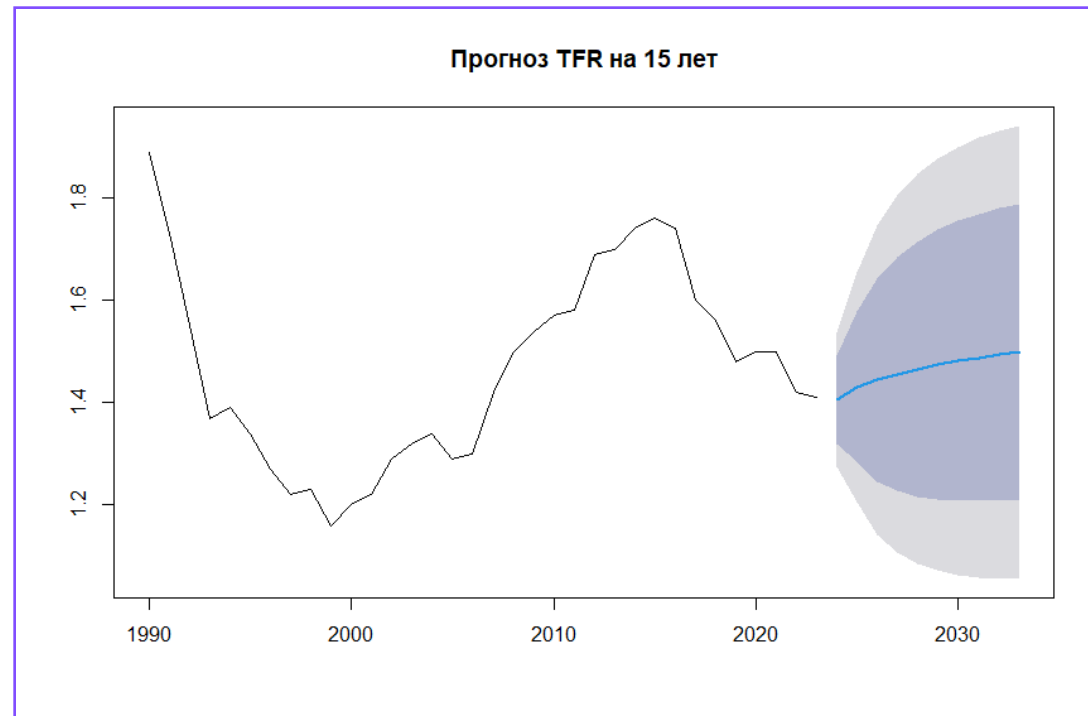
Но моделью в целях прогнозирования надо пользоваться осторожно, учитывая особенности ряда, на котором строится прогноз.

Применение ARIMA: суммарный коэффициент рождаемости в России

ARIMA-модель и реальные данные СКР, 1990-2023 гг. Точная аппроксимация



ARIMA-прогноз СКР России с 2024 по 2039 гг. неудовлетворительный, т.к. широкий интервал результатов прогноза



Прогноз миграции: сложное в простом

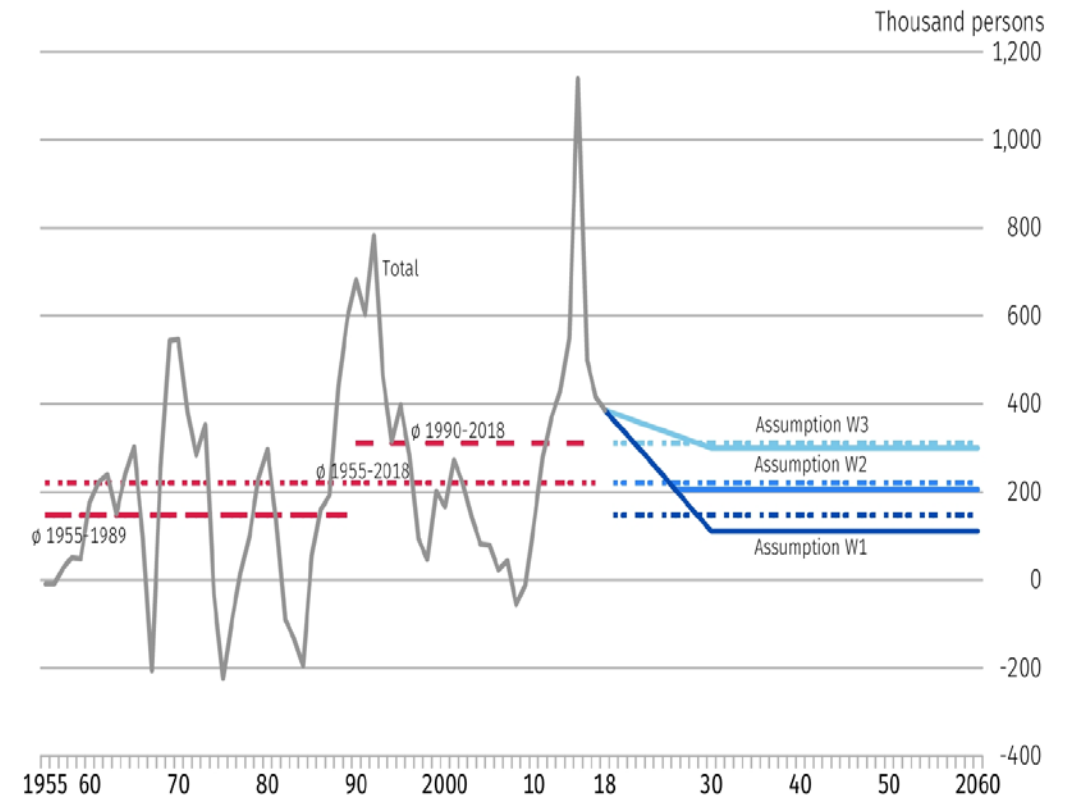
Миграция менее инерционный процесс по сравнению с процессами рождаемости и смертности. Из-за высокой чувствительности к социально-экономическим изменениям ее показатели сложно прогнозировать на длительный период.

Распространённый подход к прогнозированию международной миграции – выделение среднего миграционного прироста за определенный период и его пролонгация в будущее.

Часто миграцию увязывают с динамикой изменения населения в рабочих возрастах или прогнозами потребности в рабочей силе, или ее избытка.

Регрессионные модели в гравитационной форме применяются при моделировании внутренней миграции

Figure 19
Balance of migration across Germany's borders
From 2019, assumptions of the 14th coordinated population projection





(Б) ДЕМОГРАФИЧЕСКИЕ МОДЕЛИ

В центре внимания возраст и поколения (когорты)

Применение для оценки возрастных и поколенческих характеристик демографических процессов

Условие: известны календарные показатели уровня процессов

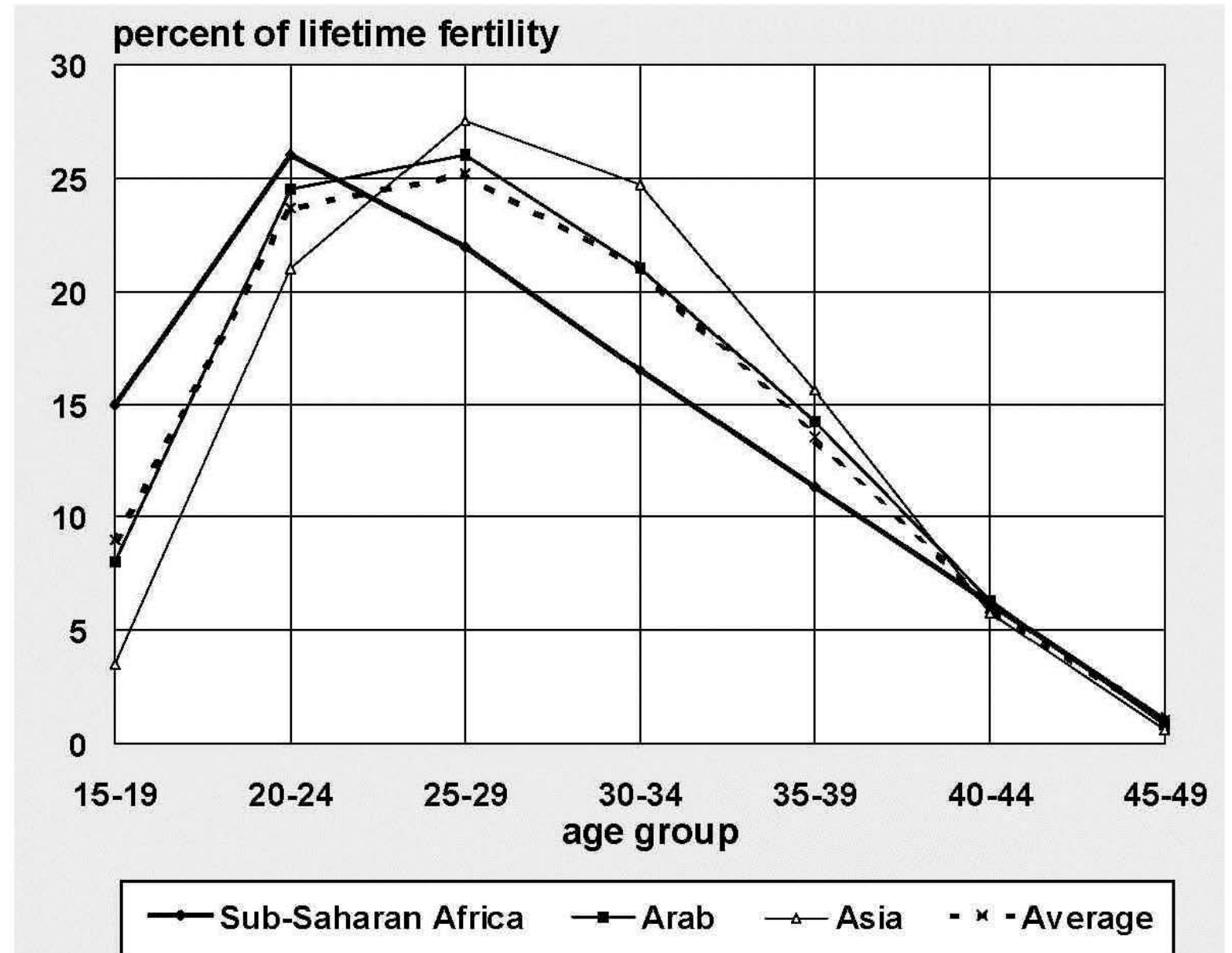
- Смертность: модельные таблицы смертности, логит-модель Брасса, модель Гомперца-Мейкема и др.
- Рождаемость: модельные таблицы рождаемости, модель Коула-Трассела, модель Брасса-Гомперца
- Миграция – модель Роджерса-Кастро.

Возможна оценка неизвестных показателей для когорт, не завершивших свою репродуктивную деятельность или жизненный путь (Модель Брасса и Браса-Гомперца).

Для подбора моделей можно воспользоваться программами Mortpak, приложениями в R, типовыми таблицами смертности в Excel Отдела народонаселения ООН и пр.

Сравнение таблиц моделей рождаемости при СКР равном 4.0

Выбор возрастной модели
рождаемости
при условии, что
суммарный суммарного
коэффициент
Рождаемости (СКР) равен
4 рождения на одну
женщину



Замечание:

- Прогнозы рождаемости, выполненные на основе показателей условных поколений обязательно должны согласовываться с тенденциями рождаемости в реальных поколениях и наоборот.



(В) МЕТОД АНАЛОГИЙ ИЛИ ТАРГЕТИРОВАНИЕ

Часто применяется на практике

Метод аналогий: преимущества и недостатки

- Суть метода: на примере (по аналогии) стран/регионов, близких в социально-историческом отношении, но ушедших вперед в социально-экономическом и демографическом развитии: (а) строится модель будущих изменений параметров демографических процессов, (б) задаются целевые (конечные) показатели на финальную дату прогноза.
- Метод особенно часто применим в случаях, когда нет долгосрочных или явных трендов (крайний случай – отсутствуют данные)
- Часто применяется в региональных прогнозах, задаются цели сходимости, как правило, к передовым или средним показателям (таргетирование).
- Как и в случае трендовых моделей, метод теряет эффективность, когда резко меняются исторические условия (новые технологии, изменения в составе населения, социальные преобразования и пр.)

Примеры использования метода аналогий:

- Прогноз изменений в продолжительности жизни и возрастных характеристиках смертности для России в 2000-2010-х гг. разрабатывался в НИУ ВШЭ по эмпирической модели изменений в странах Восточной Европы.
- Прогноз продолжительности жизни для штатов США в 2000-х гг. выполнялся как «догоняющий» до штата с самой высокой продолжительностью жизни (Гавайи);
- Прогнозный уровень смертности в старших группах может задаваться в прогнозах смертностью младших в предшествующем временном периоде.
- Модель изменения рождаемости сельского населения может разрабатываться на основе городского.

Модели роста продолжительности жизни ООН (лет)

Начальная продолжительность при рождении	мужчины					женщины				
	Очень быстрый рост	Быстрый рост	Средний рост	Медленный рост	Очень медленный рост	Очень быстрый рост	Быстрый рост	Средний рост	Медленный рост	Очень медленный рост
40,0-42,5	2,5	2,1	1,9	1,3	1,1	2,6	2,3	2,0	1,4	1,1
42,5-45,0	2,8	2,4	2,0	1,4	1,1	3,0	2,5	2,1	1,5	1,2
45,0-47,5	3,0	2,5	2,1	1,8	1,2	3,1	2,6	2,2	1,9	1,3
47,5-50,0	3,0	2,6	2,2	1,8	1,3	3,2	2,7	2,3	1,9	1,4
50,0-52,5	3,2	2,7	2,3	1,9	1,4	3,4	2,9	2,4	2,0	1,5
52,5-55,0	3,6	2,7	2,4	2,0	1,5	3,7	3,0	2,6	2,0	1,7
55,0-57,5	3,7	2,6	2,4	2,0	1,5	3,7	3,0	2,6	2,0	1,8
57,5-60,0	3,8	2,6	2,4	2,0	1,5	4,0	3,0	2,6	2,0	1,8
60,0-62,5	3,4	2,5	2,2	1,7	1,0	3,8	3,0	2,6	2,0	1,7
62,5-65,0	3,2	2,3	1,9	1,5	0,9	3,6	2,8	2,4	2,0	1,5
65,0-67,5	3,2	2,0	1,6	1,0	0,7	3,5	2,6	2,3	1,8	1,0
67,5-70,0	2,0	1,5	1,2	1,0	0,6	3,3	2,6	2,1	1,5	1,0
70,0-72,5	1,5	1,2	1,0	0,8	0,5	3,0	2,0	1,8	1,2	0,8
72,5-75,0	1,3	1,0	0,9	0,8	0,5	2,0	1,5	1,2	0,9	0,8
75,0-77,5	1,1	0,8	0,6	0,5	0,5	1,8	1,2	1,0	0,8	0,7
77,5-80,0	1,0	0,5	0,5	0,4	0,4	1,6	1,0	0,9	0,7	0,5
80,0-82,5	0,9	0,5	0,5	0,4	0,4	1,4	0,8	0,6	0,5	0,5
82,5-85,0	0,8	0,5	0,5	0,4	0,3	1,3	0,5	0,5	0,4	0,4
85,0-87,5	0,7	0,5	0,4	0,3	0,2	1,3	0,5	0,4	0,3	0,2
87,5-90,0	0,6	0,5	0,4	0,3	0,2	1,2	0,5	0,4	0,3	0,2
90,0-92,5	0,6	0,5	0,4	0,3	0,2	0,8	0,5	0,4	0,3	0,2

Это результат обобщения мирового опыта увеличения продолжительности жизни.

Темп роста зависит от ее величины.



(Г) СТРУКТУРНЫЕ МОДЕЛИ

Структурные модели – широкий класс моделей демографических изменений, их механизмов и факторов

Структурные модели или тренды?

	Структурные (факторные) модели типа $Y=f(x)$, где X -факторы	Трендовые модели
Достоинства	Отражают причинно-следственный механизм и факторы, определяющие изменение прогнозируемого демографического показателя.	Не вдаваясь в причины процессов, предсказывают будущее на основе анализа временного ряда интересующего демографического показателя.
Недостатки	Из-за сложности и изменчивости демографических процессов, как правило, не обладают динамическими свойствами.	Не учитывают будущие изменения в трендах

Пример: Модель Андриенко-Гуриева (2006)

внутренних и внешних миграционных потоков

$$M^t[ij] = k \cdot \frac{(M^{t-1}[ij])^\alpha \cdot (M^{t-1}[ji])^\beta}{d[ij]^\rho} \cdot \prod_{k=1}^c (X^{t-1}[j])^{\sigma_k} \prod_{l=1}^d (X^{t-1}[i])^{\delta_l}$$

- **Независимая переменная:** количество мигрантов из региона А в регион В;
- **Зависимые переменные**
 - Расстояние;
 - Уровень жизни (доходы и занятость)
 - Демографические (численность населения, детских и трудоспособных групп, коэффициенты рождаемости и смертности, ОПЖ)
 - Предоставления общественных благ (телефоны, врачи и больничные койки)

- - характеристики динамики (число мигрантов из А в В за прошлые годы);
- - промежуточные возможности (средневзвешенная характеристик по другим регионам в весах, равному количеству мигрантов из А в другие регионы);
- - макроэкономические переменные (рост ВВП, экспорт)



(Д) МЕТОД ОПРОСА ЭКСПЕРТОВ

Следует проводить ежегодно, если есть настоящие эксперты

Применение экспертных опросов

- Опросы экспертов применяются при разработке как детерминистских, так и вероятностных демографических прогнозов.
- Опросы населения в сфере рождаемости – выявление репродуктивных установок. Полученные результаты эффективны в краткосрочном периоде для уточнения рождаемости в когортах, у которых не завершился репродуктивный период.



РЕЗУЛЬТАТЫ ОПРОСОВ ЭКСПЕРТОВ В НИУ ВШЭ

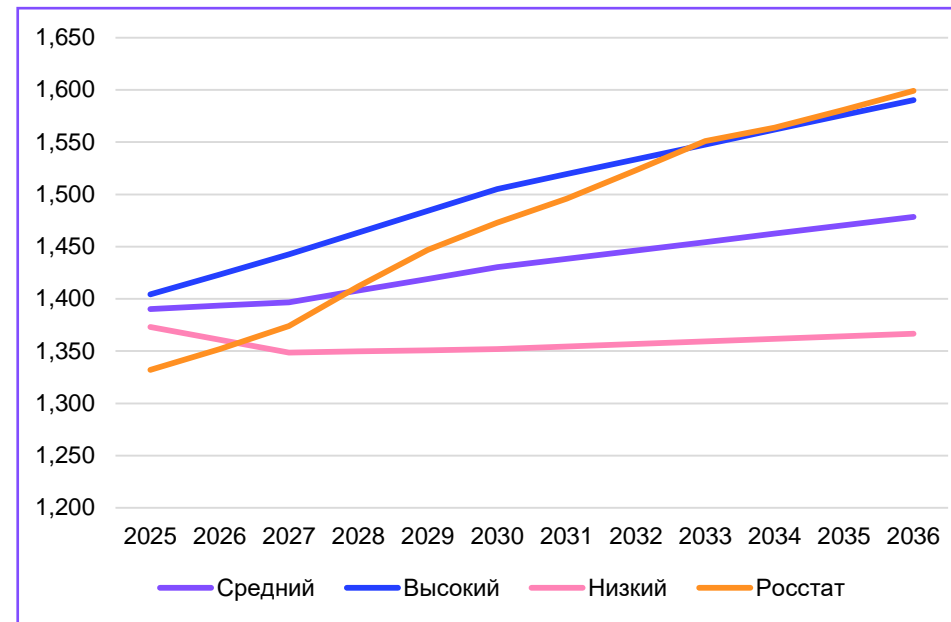
40 экспертов; увязка с возможными сценариями социально-экономического и геополитического развития



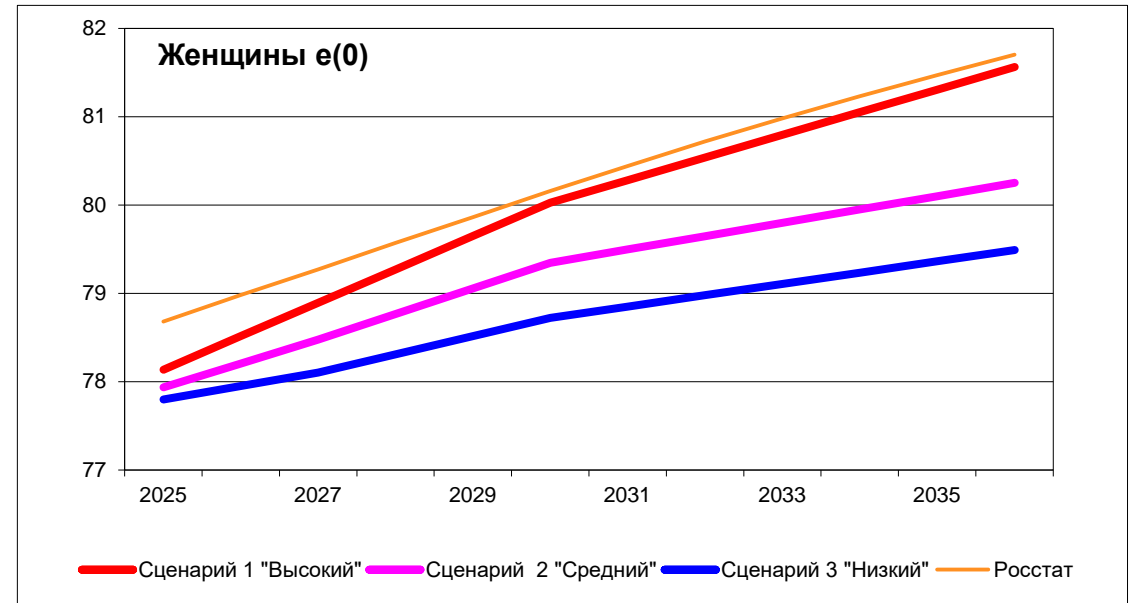
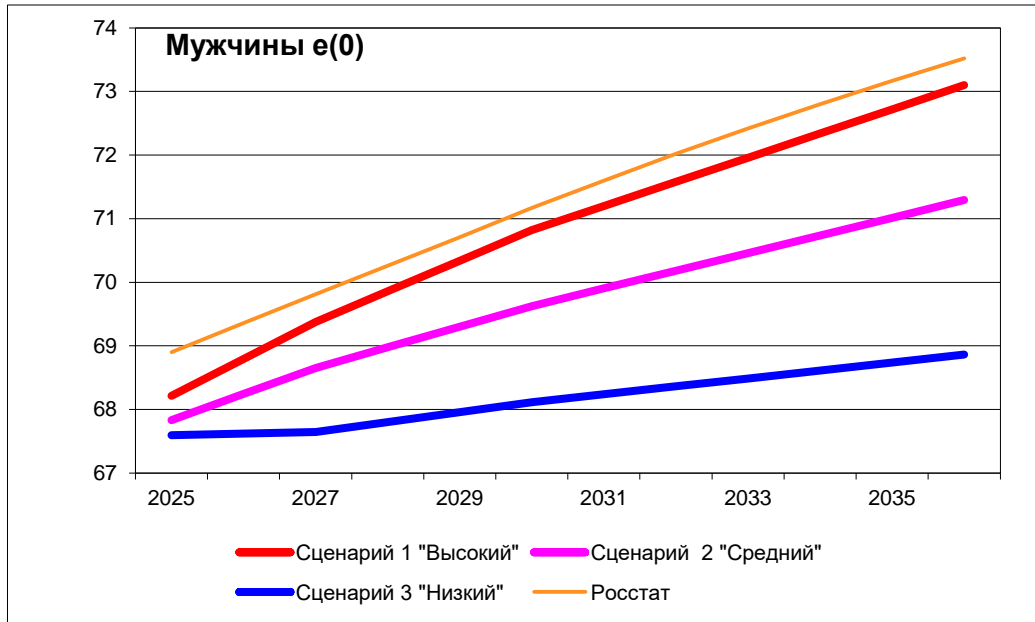
Суммарный коэффициент рождаемости (рождений на одну женщину) (средняя арифметическая ответов)

	Опрос	2025	2027	2030	2036*
Сценарий 1 Высокий	2024	1,42	1,45	1,51	1,56
	2025	1,40	1,44	1,51	1,59
Сценарий 2 Средний	2024	1,40	1,41	1,43	1,48
	2025	1,39	1,40	1,43	1,48
Сценарий 3 Низкий	2024	1,38	1,36	1,35	1,35
	2025	1,37	1,35	1,35	1,37
Росстат (средний)	2023	1.332	1.374	1.473	1.599

* по опросу 2024 г. оценка приводится для 2035г.;



Ожидаемая продолжительность жизни (ОПЖ) при рождении в 2025-2036 гг., лет



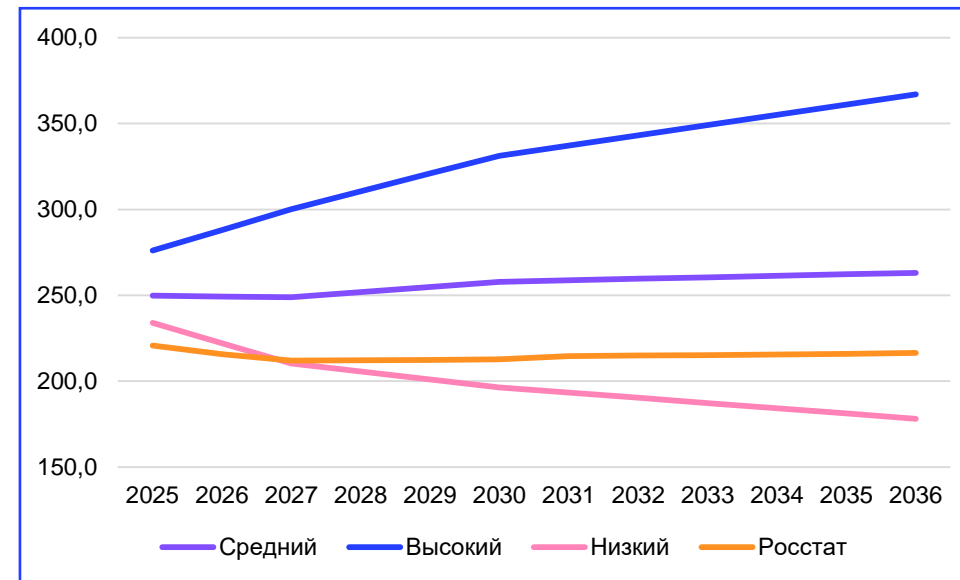
Оценки получены методом экстраполяции средних значений для 2025, 2027, 2030 и 2036 гг.



Миграционный прирост (тыс. человек)

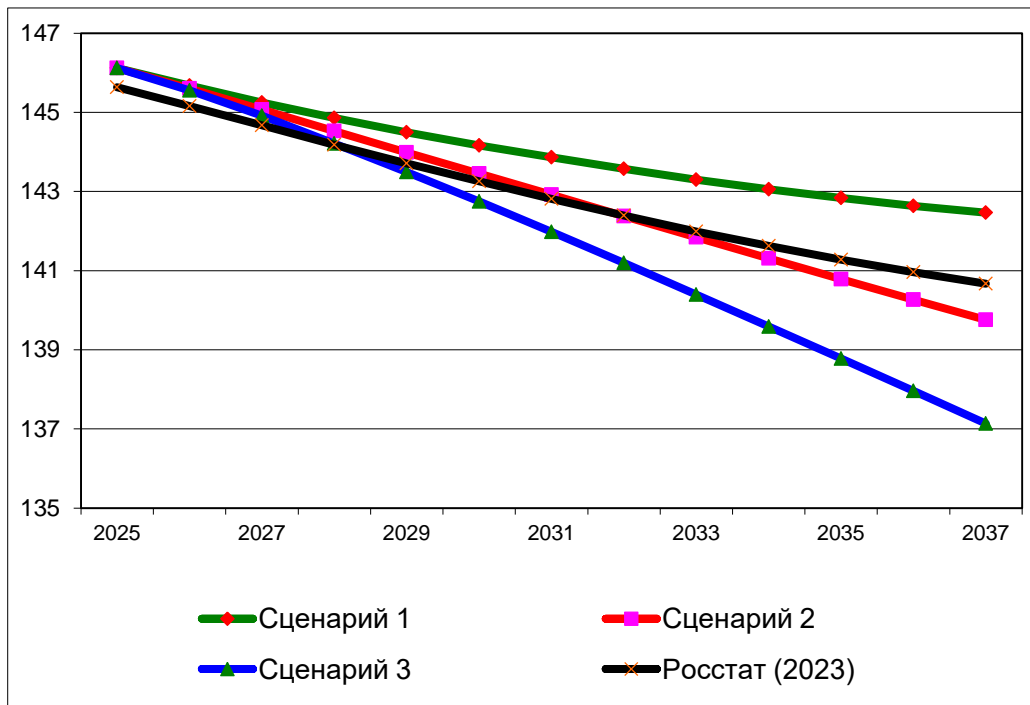
Сценарии	Опрос	2025	2027	2030	2036
Сценарий 1 Высокий	2025	276,0	300,1	331,3	366,9
	2024	189,3	216,4	241,6	270,3
сценарий 2 Средний	2025	249,8	248,8	257,9	263,1
	2024	168,9	181,3	202,4	220,4
сценарий 3 Низкий	2025	234,0	210,3	196,4	178,2
	2024	152,2	148,9	143,4	139,8
Росстат (средний)	2023	220,7	212,1	212,7	216,4

*расчетная величина





Численность населения по 4 сценариям на начало года 2025-2037 гг. (млн человек)



Годы	Сценарии			Росстат (средний)
	1 (высокий)	2 (средний)	3 (низкий)	
2025	146.1	146.1	146.1	145.6
2026	145.7	145.6	145.5	145.2
2027	145.2	145.1	144.9	144.7
2028	144.9	144.5	144.2	144.2
2029	144.5	144	143.5	143.7
2030	144.2	143.4	142.7	143.3
2031	143.9	142.9	142	142.8
2032	143.6	142.4	141.2	142.4
2033	143.3	141.8	140.4	142.0
2034	143.1	141.3	139.6	141.6
2035	142.8	140.8	138.8	141.3
2036	142.6	140.3	138	141.0
2037	142.5	139.8	137.1	140.7

Д) ТЕОРЕТИЧЕСКИЕ ОСНОВЫ РАЗРАБОТКИ ДЕМОГРАФИЧЕСКИХ ПРОГНОЗОВ

Теория демографического перехода остается

Теоретические основы разработки прогнозных сценариев

Фундамент

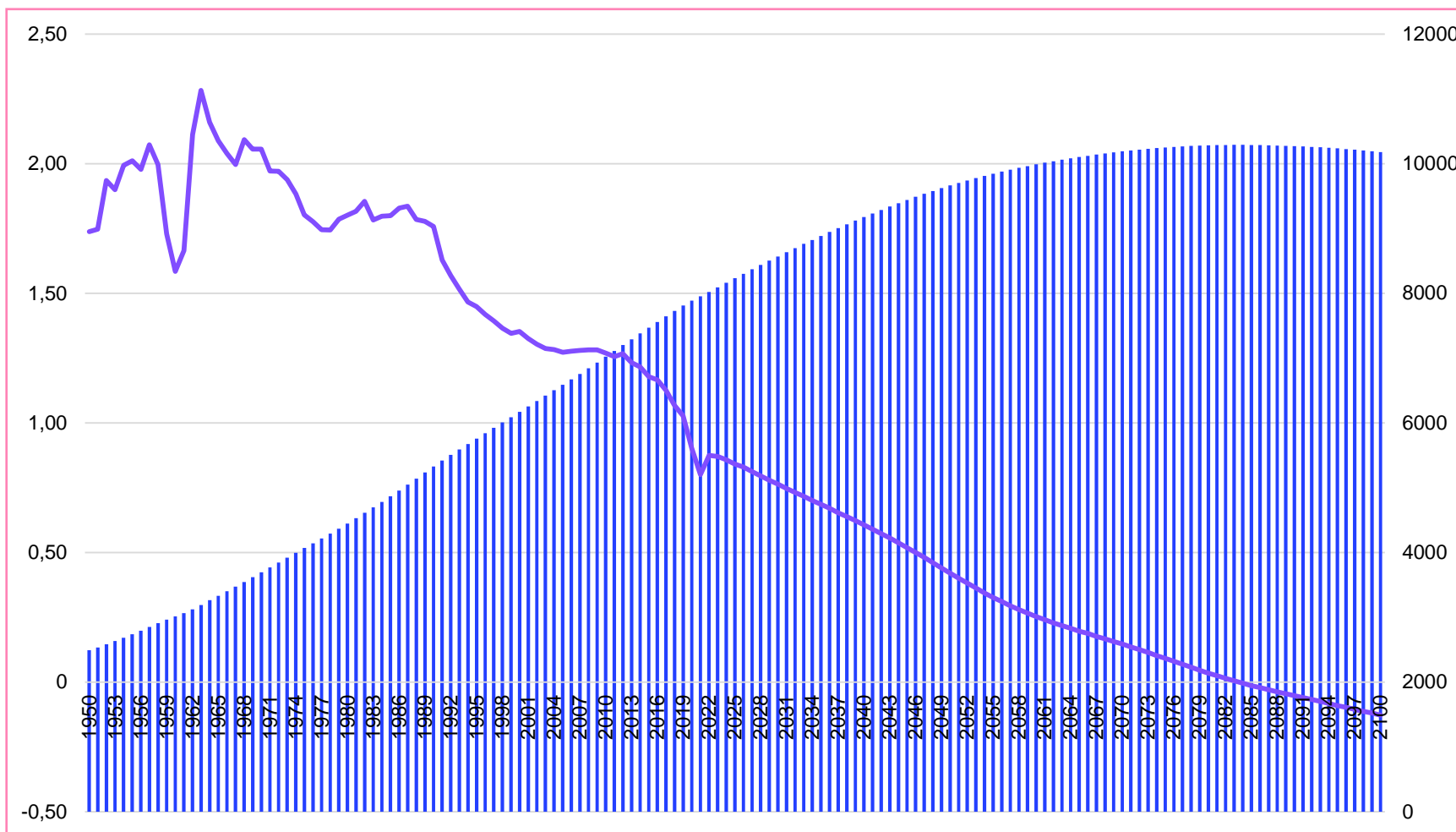
Теория демографического перехода в современном виде:

- отказ от идеи установления нового равновесия между рождаемостью и смертностью, обеспечивающего простое воспроизводство, в конце демографического перехода
- Допускается, что рождаемость не только на уровне стран, но и на глобальном уровне в конце 21 века будет ниже 2 рождений на женщину.
- Допускается депопуляция в глобальном масштабе в конце 21 века, а не стабилизация численности мирового населения, как предполагалось ранее.

Теоретические «надстройки» на теорией Дем. перехода

- Для рождаемости: теория «второго демографического перехода», экономические модели рождаемости и др.
- Для смертности: теория эпидемиологического перехода, теория эпидемиологической революции, модели инвестиций в здравоохранение и др.
- Для миграции: гравитационные модели, экономические теории миграции, теория миграционного перехода и др.

С 2024 по 2100 Мировое население увеличится с 8,2 до 10,2 млрд, но темпы его прироста заметно уменьшатся с 0.87% до -0,13% в 2100 г.



Источник: UN Population Division, World Population Prospects: The 2024 Revision

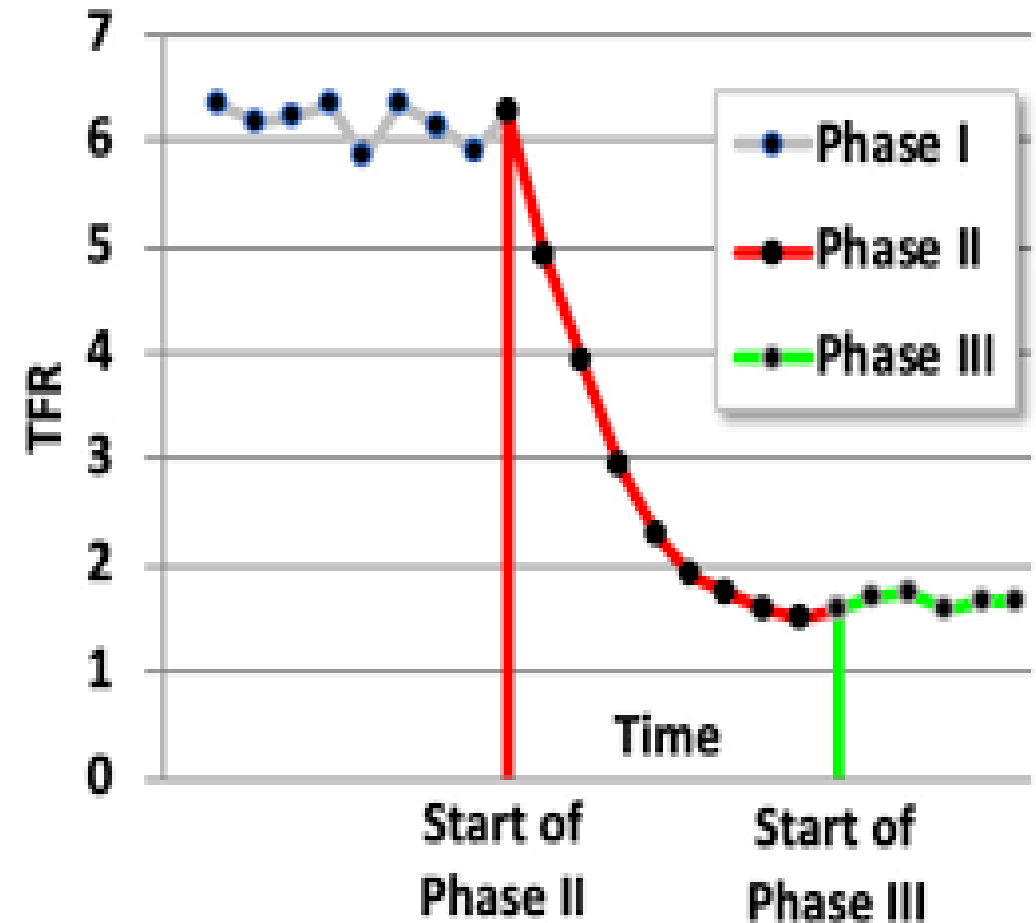
Динамика численности основных регионов мира, 1500-2050 гг.

Регион	1500	1600	1700	1750	1800	1850	1900	1950	2000	2050
Численность населения (млн. человек)										
Мир	458	580	682	791	978	1262	1650	2494	6172	9665
Африка	86	114	106	106	107	111	133	228	831	2467
Азия	243	339	436	502	635	809	947	1368	3748	5280
Европа	84	111	125	163	203	276	408	549	728	703
Лат. Америка	39	10	10	16	24	38	74	168	521	730
Сев. Америка	3	3	2	2	7	26	82	168	313	427
Океания	3	3	3	2	2	2	6	13	31	58
Доля в населении Земного шара (в %)										
Африка	18.8	19.7	15.5	13.4	10.9	8.8	8.1	9.1	13.5	25.5
Азия	53.1	58.4	63.9	63.5	64.9	64.1	57.4	54.9	60.7	54.6
Европа	18.3	19.1	18.3	20.6	20.8	21.9	24.7	22.0	11.8	7.3
Лат. Америка	8.5	1.7	1.5	2.0	2.5	3.0	4.5	6.7	8.4	7.6
Сев. Америка	0.7	0.5	0.3	0.3	0.7	2.1	5.0	6.7	5.1	4.4
Океания	0.7	0.5	0.4	0.3	0.2	0.2	0.4	0.5	0.5	0.6

Источник: UN Population Division, World Population Prospects: The 2024 Revision

Пример: Стадии снижения (перехода) суммарного коэффициента рождаемости (число рождений на одну женщину)

- Стадия 1: Высокая рождаемость до начала демографического перехода не моделируется.
- Стадия 2: Снижение рождаемости (переход). Моделируется двойной логистической функцией с применением байесовской иерархической модели.
- Стадия 3: Низкая рождаемость. Используется авторегрессия первого порядка $AR(1)$ с применением байесовской иерархической модели.
- Эта модель положена в основу прогноза рождаемости в ООН
- Источник: Alkema, L., and others (2011). Probabilistic projections of the total fertility rate for all countries. *Demography*, vol. 48, No. 3, pp. 815-39.



Байесовский подход и его использование л построения вероятностных прогнозов населения стран мира в Отделе народонаселения ООН

Содержание

Вероятности $P(A_k)$ называют априорными, т. е. до выполнения опыта

Условные вероятности этих событий $P(A_k/B)$ называют апостериорными, т.е. уточненными в результате опыта, исходом которого послужило появление события В.

В отличие от классического подхода к оценке вероятностей для байесовского подхода требуется определение априорных вероятностей для самой модели, помимо ее спецификации. Это позволяет оценивать неопределенность в отношении уникальных событий (параметры – случайные величины) и обновлять информацию при появлении нового априорного знания.

Формула Байеса

Формула полной вероятности:

- $$P(B) = \sum_1^{\infty} P(A_k) \cdot P(B/A_k)$$

Формула Байеса

- $$P(A_k/B) = \frac{P(A_k) \cdot P(B/A_k)}{P(B)}$$

Где A_k – попарно несовместимые события, определяют полное пространство событий

Байесовский подход : достоинства и недостатки

Достоинства

- Построение сложных иерархических моделей (например: страны-регионы-мир), учитывающие структуру имеющихся данных.
- Привлечение разнообразной обновляемой информации.
- Работа с малыми выборками
- Гибкость модели.
- Широкий круг приложений.

Недостатки

- Выбор априорных вероятностей (распределений) субъективен
- Вычислительная сложность

Пример: Байесовский прогноз рождаемости ООН

- Априорная информация:
- Предположение о наиболее вероятном пути снижения рождаемости в стране, которая еще не завершила демографический переход.
- Анализируются **исторические траектории рождаемости сотен стран**, которые уже прошли путь от высокой к низкой рождаемости. Эти траектории группируются в паттерны (например, "быстрое снижение", "медленное снижение", "зависание на среднем уровне").
- Для страны, находящейся на СКР=4.0, априорным распределением будет набор вероятных будущих траекторий, основанных на опыте других стран, которые были на этом уровне.
- **Априорное распределение для темпа изменения СКР**
- **Априорное распределение для "нижнего предела" рождаемости.** На основе данных по странам с самой низкой рождаемостью (например, Южная Корея, Сингапур, страны Южной Европы). Априорно считается, что СКР из опыта вряд ли надолго опустится ниже ~0.7 (экстремальный минимум) и с большей вероятностью будет колебаться вокруг 1.5-1.9 в долгосрочной перспективе с возможным последующим слабым ростом
- **Априорные корреляции между демографическими параметрами**

Работа модели:

- **Шаг 1 (Априорная информация по миру):** Для страны X, в которой снижается рождаемость, можно получить априорное распределение СКР в будущем: Из опыта всех стран мира следует, что в странах с СКР=5.0 и ИЧР=0.55 рождаемость снижалась со скоростью Y.
- **Шаг 2 (Данные по стране):** Затем модель учитывает имеющиеся **фактические данные** по стране X за последние 20 лет. Если страна X снижала рождаемость *быстрее*, чем типичная априорная траектория, модель корректирует прогноз в сторону более быстрого снижения СКР.
- **Шаг 3 (Апостериорный прогноз):** Итоговый прогноз (апостериорное распределение) — это "усреднение" между общемировым априорным паттерном и конкретными данными страны. **Чем качественнее и полнее данные по стране, тем больший вес они имеют по сравнению с априорным распределением.**

Пример: Байесовский прогноз рождаемости ООН

- Для моделирования снижения рождаемости во Фазе II используется двойная логистическая функция со следующими характеристиками:
- Во-первых, это стохастическая функция (гетероскедастичный член ошибки в уравнении).
- Во-вторых, параметры функции выбираются из небольшого числа возможностей.
- В-третьих, моделируются значения параметра для разных стран как взятые из «глобального» распределения. Это приводит к оценкам, которые учитывают информацию по другим странам, что делает модель иерархической.
- Последнее важно, поскольку для отдельной страны данных мало (для большинства стран это не более 12 временных периодов), и оценка специфичной для страны двойной логистической кривой может быть неустойчивой, так как включает оценку пяти параметров по 12 или менее точкам данных. Иерархическая модель стабилизирует эту оценку.

Прогноз для стран на второй стадии снижения рождаемости

- Моделирование снижения СКР с помощью двойной логистической функции

- Результат $f_{c,t+1} = f_{c,t} - g(f_{c,t}/\theta_c) + \alpha_{c,t}$,

- где
$$g\left(\frac{f_{c,t}}{\theta_c}\right) = \frac{d_c}{1 + \exp\left(\frac{2\ln 9}{\Delta_{c1}}(f_{c,t} - \sum_i \Delta_{ci} + 0.5\Delta_{c1})\right)} + \frac{d_c}{1 + \exp\left(\frac{-2\ln 9}{\Delta_{c3}}(f_{c,t} - \Delta_{c4} - 0.5\Delta_{c3})\right)}$$

- Где $f_{c,t}$ – СКР в стране с в момент t

- $\theta_c = (\Delta_{c1}, \Delta_{c2}, \Delta_{c3}, \Delta_{c4}, d_c)$

d_c – максимум сокращения СКР (достигается в точке перегиба между двумя логистическими кривыми)

Δ_{c1} — это интервал, на котором растет сокращение СКР от $0.1 d_c$ до $0.8 d_c$.

Δ_{c2} — это "плато" высокой скорости снижения СКР (от $0.8 d_c$ до d_c).

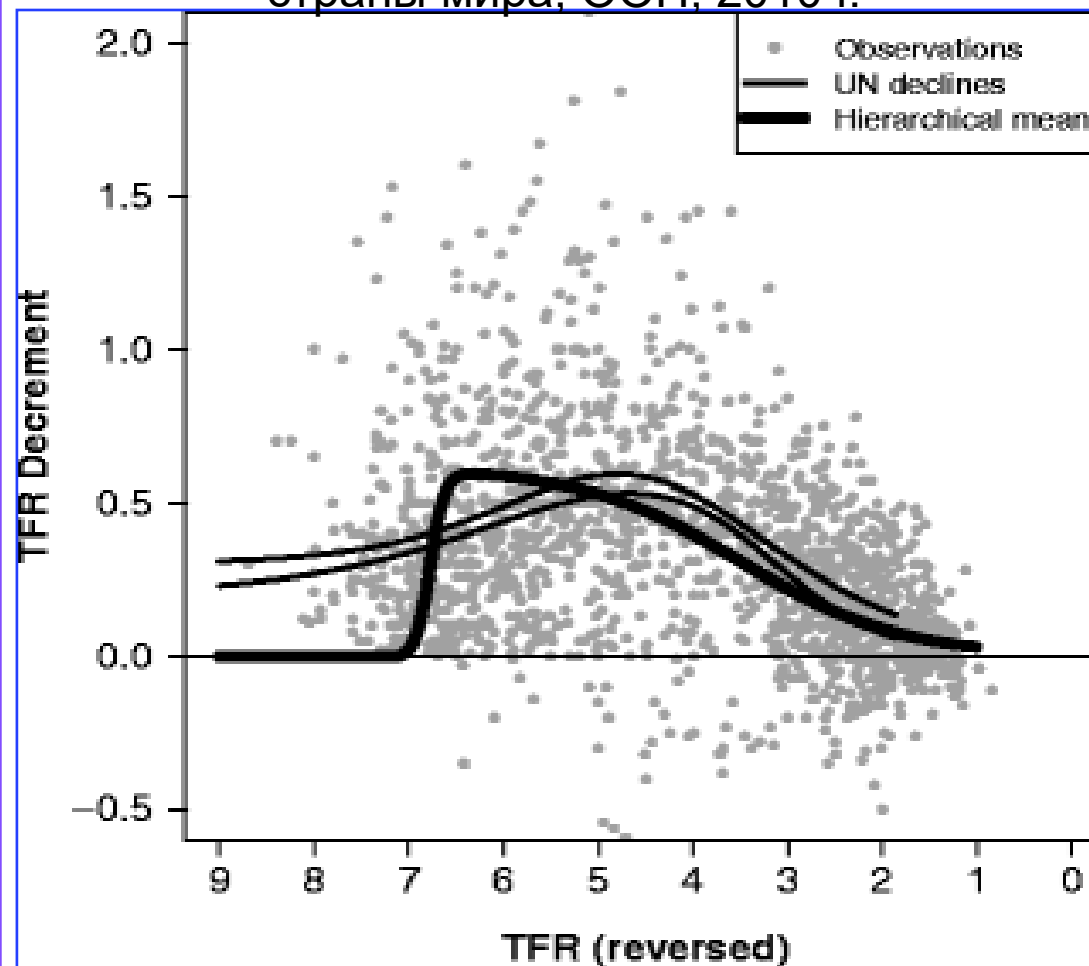
Δ_{c3} — это интервал, где скорость снижения падает от $\sim 0.8 d_c$ до $\sim 0.1 d_c$.

Δ_{c4} -это интервал, на котором снижение затухает (убыль $\sim 10\%$ от d_c)

θ_c описывает историю снижения рождаемости в стране с

разными уровнями СКР в начале, в максимуме и в конце стадии 2.

Зависимость сокращения СКР от его величины, страны мира, ООН, 2010 г.



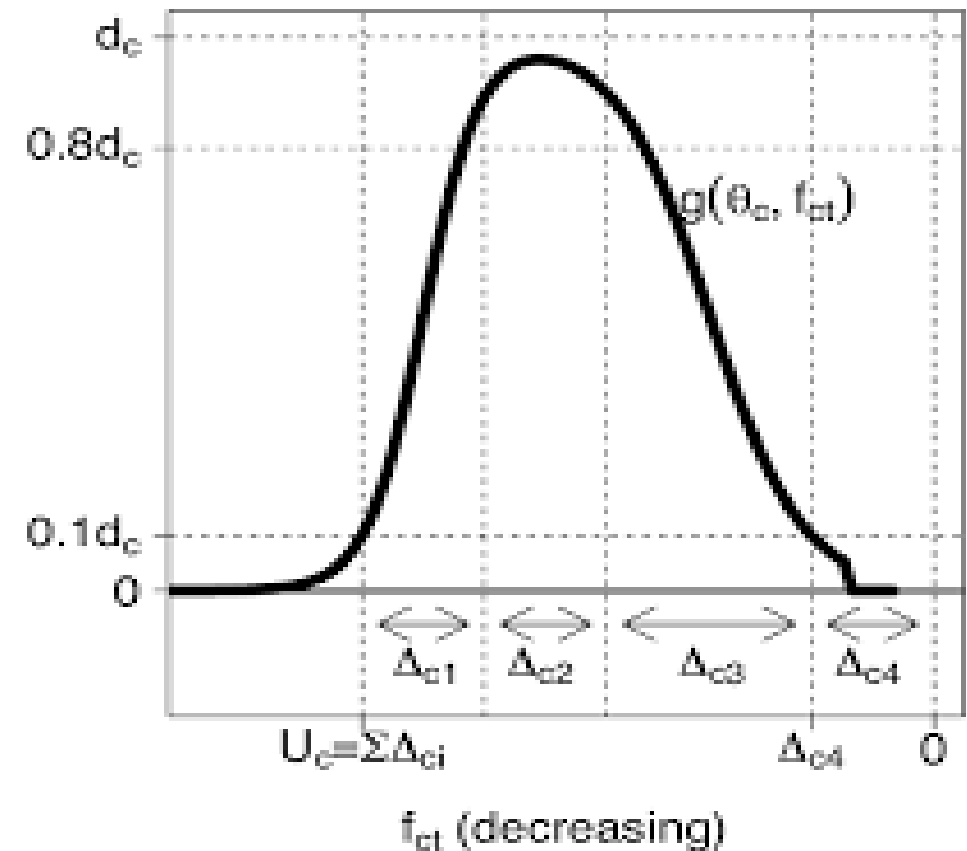
Продолжение

- $\alpha_{c,t} \sim N\left(0, \sigma(t, f_{c,t})^2\right)$

Где $\sigma(t, f_{c,t})^2$ - функция, которая описывает, как стандартное отклонение ошибки меняется в зависимости от уровня рождаемости и времени.

- Специфические параметры $\theta_c = (\Delta_{c1}, \Delta_{c2}, \Delta_{c3}, \Delta_{c4}, d_c)$ для каждой страны (с) берутся из глобального (мирового) распределения, параметры (или гиперпараметры) ϕ которого, в свою очередь, сами имеют диффузное априорное распределение. Эти параметры оценены по данным стран, которые прошли стадию 2 снижения рождаемости.
- Полученная байесовская иерархическая модель оценивается с использованием метода Монте-Карло по схеме марковской цепи (MCMC).

Параметры Δ , отражающие снижение СКР



Резюме байесовского подхода к прогнозированию рождаемости во всех странах мира

- Оценка параметров $\theta_c = (\Delta_{c1}, \Delta_{c2}, \Delta_{c3}, \Delta_{c4}, dc)$ с малым числом наблюдений дается в рамках байесовской иерархической модели, которая позволяет использовать для оценки данные по всем странам мира
- Байесовская иерархическая модель предполагает, что для всех стран каждый из неизвестных параметров снижения СКР извлекается из распределения вероятностей, которое представляет диапазон возможных значений этого параметра по всем странам.
- Для каждого параметра в конкретной стране это распределение вероятностей, основанное на общемировом опыте, затем обновляется с помощью теоремы Байеса с учётом наблюдаемого снижения в данной стране. В результате получается апостериорное (послеопытное) распределение интересующего параметра.
- Полученные оценки (выборки из апостериорного распределения) можно рассматривать как взвешенные средние значения между «общемировой картиной» и информацией из данных конкретной страны

Глобальный
уровень: все
страны мира

Формула Байеса
$$f(\theta | y) = \frac{F(y|\theta)\pi(\theta)}{f(y)}$$

Априорные
распределения
параметров

Уровень
страны С:
данные

Оценка
параметров
для страны С

Прогноз рождаемости для стран, в которых переход в рождаемости завершился

- Для прогнозирования используют модель авторегрессии первого порядка AR(1) с определенным долгосрочным (на конец периода) средним μ (до недавнего времени принимался равным 2.1 рождений на одну женщину) для всех стран, находящихся в Фазе III. Сейчас вводятся индивидуальные значения μ для стран:
- $f_{c,t+1} - \mu = \rho(f_{c,t} - \mu) + \beta_{c,t}$
- $\beta_{c,t} \sim N(0, \sigma_\beta^2)$
- где параметры ρ and σ оценивались методом максимального правдоподобия (21 страна, 54 периода) .

Дискуссии вокруг использования Байесовского подхода в демографии

- Байесовский подход – это макропрогноз для замкнутой территориальной или иной социальной системы. Ее характеристики – основа априорной информации.
- Субъективизм априорных распределений для стадии 2. Насколько верно моделировать снижение рождаемости или смертности, опираясь на коллекцию глобальных данных? Процессы в прошлом развивались в условиях, отличных от настоящего и тем более от будущего. Данные по странам нельзя объединять в одном «котле», поскольку это возвращает к раскритикованной концепции единого для всех «демографического перехода».
- Субъективный выбор показателей для конца прогнозируемого периода на стадии 3 вызывает вопросы. Эти оценки должны стать результатом прогноза.
- Насколько полученные сложным методом прогнозы точнее тех, которые получены более простыми способами? Мировой опыт используется, например, в методе аналогий с коррекцией. Модельные таблицы также учитывают мировой опыт.

Программное обеспечение для реализации байесовского метода в демографии

- **Открытые R-пакеты**, разработанные организацией Bayesian Demography Limited. Они созданы специально для демографического анализа и реализуют описанные методы (иерархические модели, MCMC-оценку).
- **demest: Ключевой пакет** для байесовского статистического моделирования в демографии (оценка и прогнозирование).
- Этот пакет пока не размещен в основном репозитории CRAN. Для установки нужно использовать devtools для загрузки напрямую из GitHub.



ЗАМЕЧАНИЯ О ФУНКЦИОНАЛЬНЫХ ПРОГНОЗАХ

Рынок труда и образование

Виды функциональных прогнозов

Прогнозы «предложения населения» (population supply or demographic supply projection)

- Выполняются на основе демографических прогнозов.

Прогнозы «спроса на население» (population demand projections)

- Выполняются на основе экономических расчетов.
- Показывают, сколько той или иной организации нужно работников или какое количество потребителей может организация обеспечить товарами/ услугами, исходя из планов развития

Общая схема реализации функциональных прогнозов «предложения населения»

$$F(i,s,t)=P(i,s,t)*K(i,s,t)$$

где:

- $F(i,j,t)$ – численность прогнозируемой функциональной группы в возрасте i лет, s -го пола в момент t ;
- $P(i,j,t)$ – прогнозная численность возрастной группы i лет s -го пола в момент t (результат демографического прогноза) ;

В качестве $K(i,s,t)$ могут использоваться следующие функции и их комбинации

- $r(i,j,t)$ – частота функциональных событий в возрастной группе i лет s -го пола в момент t ;
- $u(i,s,t)$ – отношение численности одной функциональной группы в возрасте i лет к численности другой того же пола в момент t ;
- $c(i,s,t)$ – доля функциональной подгруппы в общей численности возрастной группы i лет s -го пола в момент t .

Пример подхода к прогнозированию «спроса» на населения:

- Производственная функция Кобба-Дугласа:

$$Y = A \cdot L^{\alpha} K^{\beta}$$

- где Y — национальный доход; A — коэффициент размерности; L и K — соответственно объемы приложенного труда и капитала; α и β — константы (коэффициенты эластичности производства по труду L и капиталу K).
- Зная прогнозные значения параметров Y , A , K можно оценить параметр L (спрос на труд или на население).

Этапы функционального прогнозирования «предложения населения»

1. Постановка задачи.
2. Сбор необходимых данных.
3. Реализация демографического прогноза.
4. Получение перспективных оценок специальных параметров $\{r(i,j,t), u(i,j,t) \text{ и } c(i,j,t)\}$ (разработка сценарных гипотез, выбор метода, реализация прогностической модели).
5. Реализация функционального прогноза – оценка численности функциональных групп.
6. Анализ результатов.



(А) ПРИМЕР: ПРОГНОЗЫ ЧИСЛЕННОСТИ И ВОЗРАСТНОГО СОСТАВА РАБОЧЕЙ СИЛЫ

Прогноз строился для России в целом и ее регионов.

Общая схема расчетов численности экономически активного населения

$$L(i,s,t) = P(i,s,t) * I(i,s,t)$$

Где

- $L(i,s,t+n)$ – численность экономически активного населения мужского ($s=1$) или женского ($s=2$) пола в возрасте i ;
- $P(i,s,t)$ – численность соответствующей возрастно-половой группы;
- $I(i,s,t)$ – коэффициенты участия в рабочей силе

Источники данных:

- Демографический прогноз (три варианта).
- Обследование рабочей силы
- База данных ОЭСР (OECD) по рынку труда.

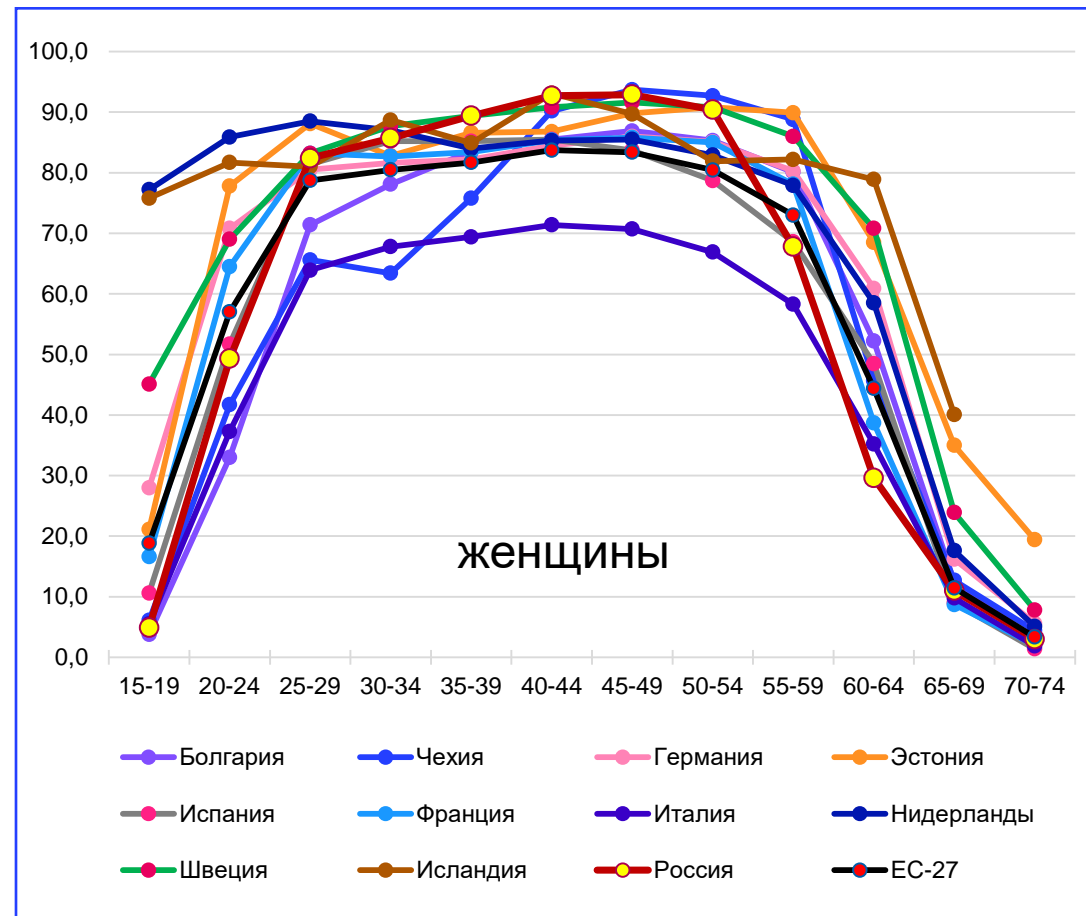
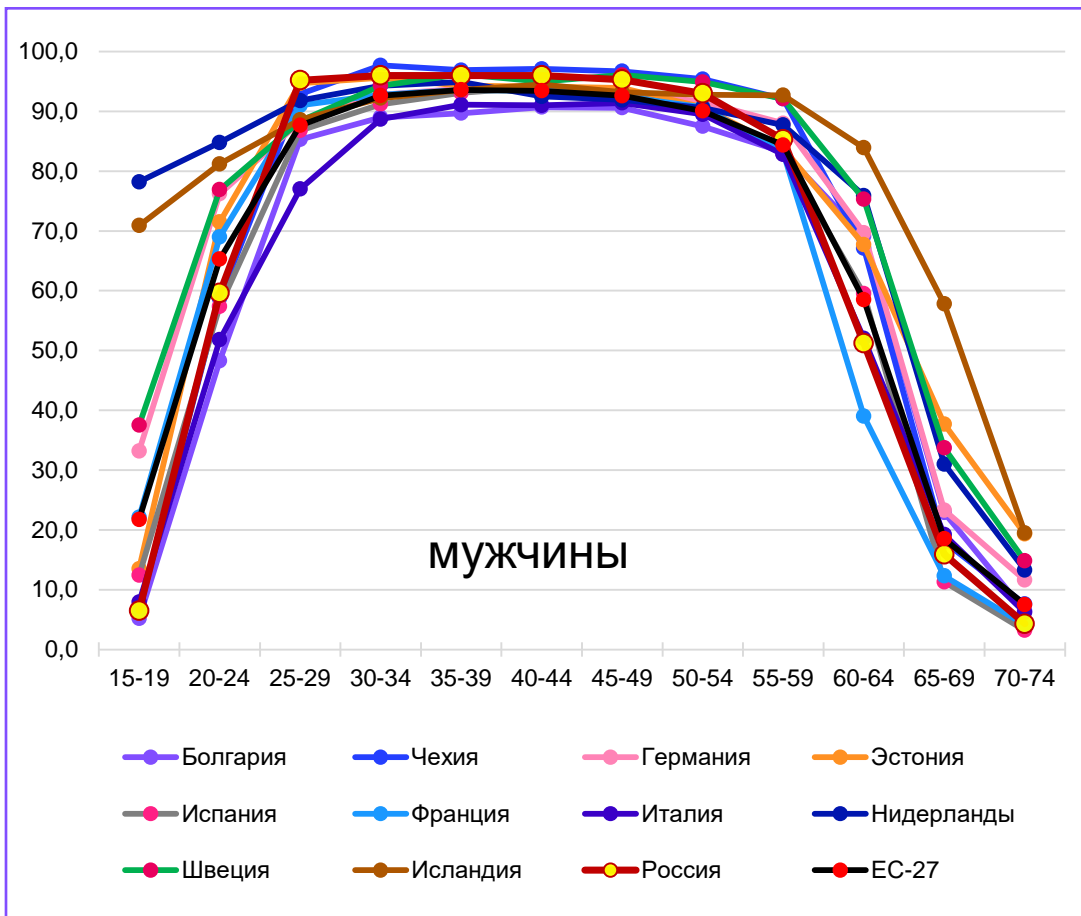
Подходы к прогнозированию коэффициентов экономической активности

- Анализ временных рядов по возрастным группам;
- Анализ временных рядов по поколениям;
- Аналоговый подход (на основе сравнительного межстранового или межрегионального анализа);
- Построение эконометрических моделей с учетом изменений в экономической конъюнктуре, состояния рынка труда, системы социального обеспечения и пр.

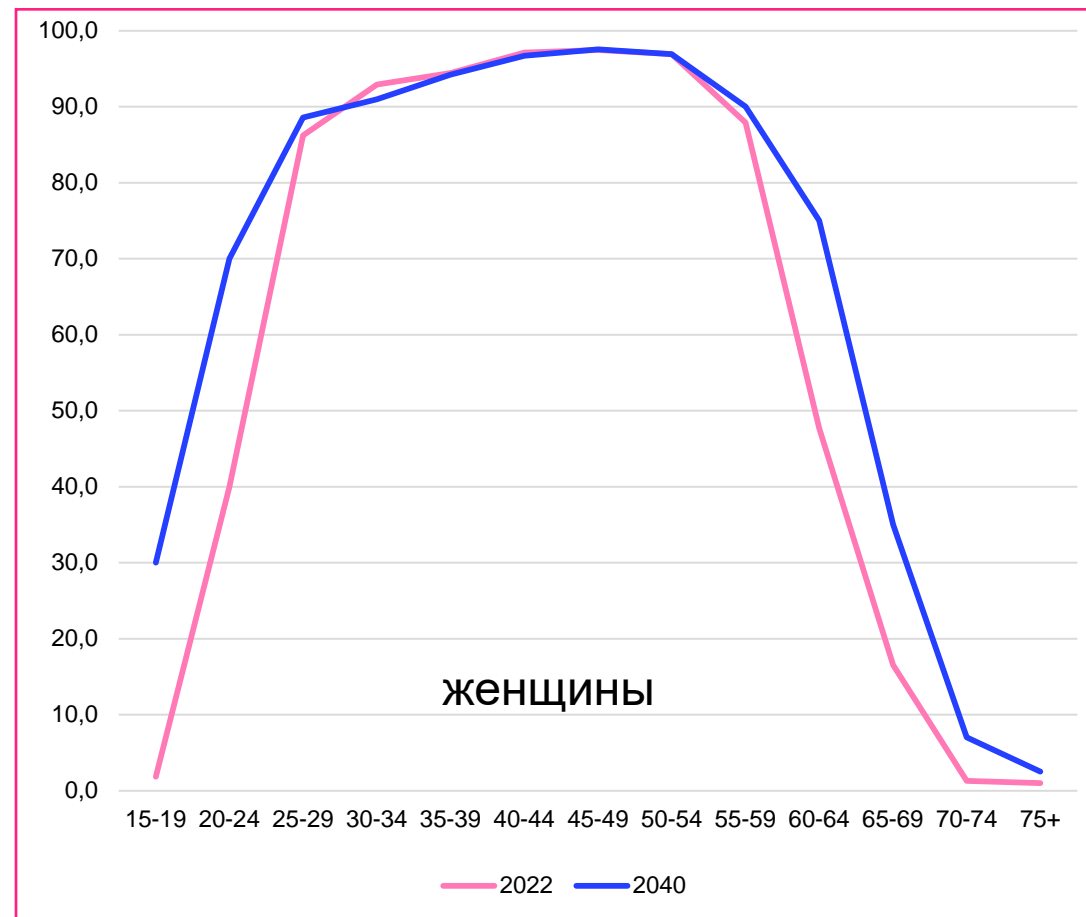
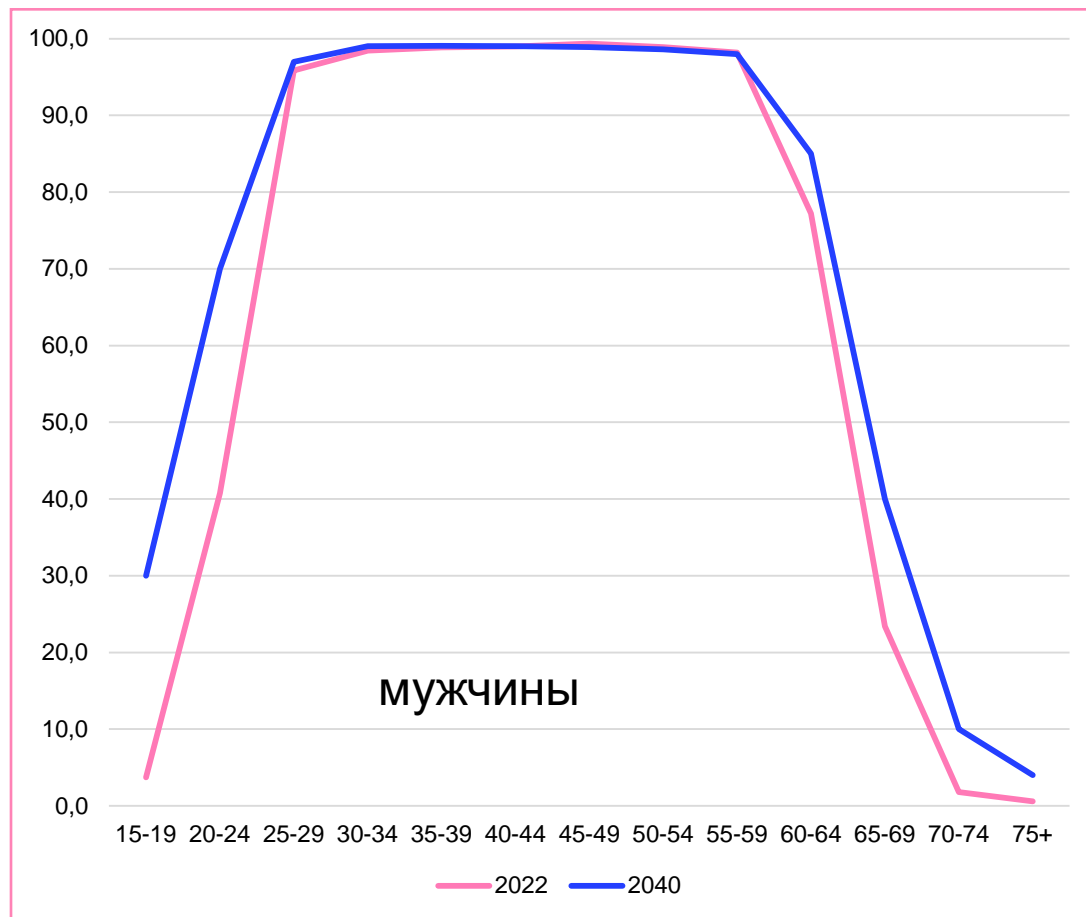
Аналоговый подход для разработки коэффициентов участия в рабочей силе

- Строилась модель изменения уровня экономической активности в старших возрастах (от 55 лет у мужчин, от 50 лет – у женщин) на основе данных для тех стран, где повышался пенсионный возраст у мужчин с 60 до 65 лет, у женщин с 55 до 60 лет.
- Допускалось повышение уровня экономической активности в молодых возрастах (до 25 лет) до средней по странам, где пенсионный возраст у мужчин составлял 65 лет, у женщин – 60 лет.
- Делалась коррекция с учетом возможных изменений в спросе на рабочую силу в условиях убыли населения в рабочих возрастах.
- В других возрастах уровень экономической активности практически не менялся, т.к. в России достигнут очень высокий уровень

Уровни участия в рабочей силе населения европейских стран и России

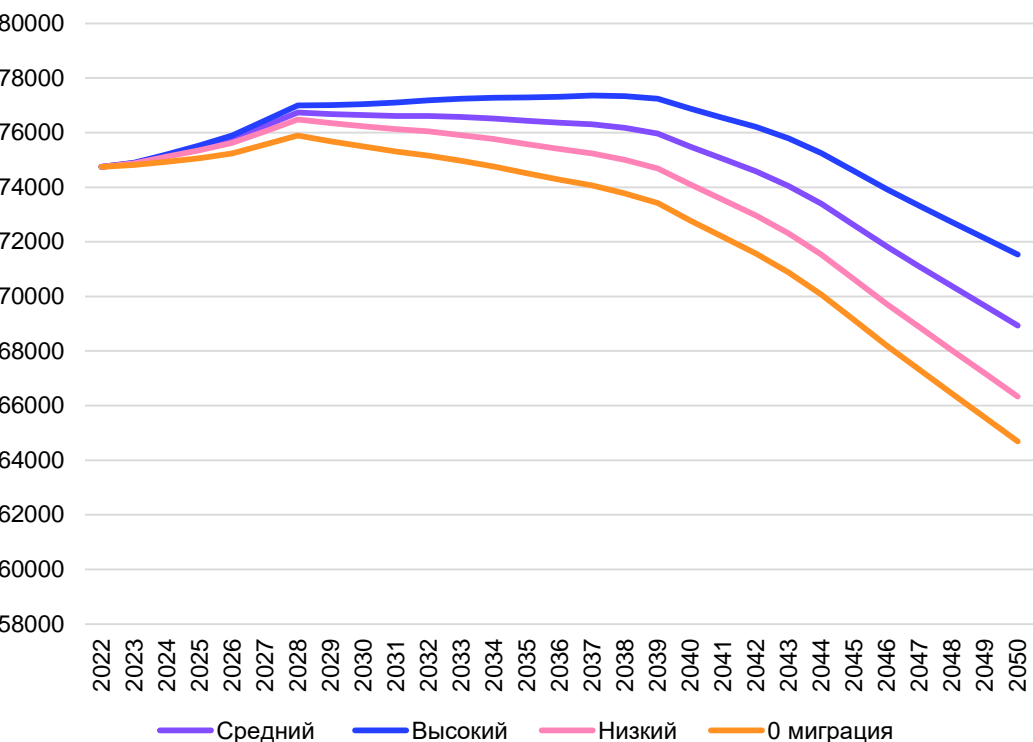


Результаты прогноза: коэффициенты участия в рабочей силе в 2022 и 2040 гг.



Результаты прогноза

Предложение рабочей силы в России, 2022–2050 гг. (в среднем за год, тыс. человек)



Численность рабочей силы, занятых и безработных (тыс. человек), оба пола, 2022–2050

Варианты	2022	2025	2030	2040	2050
Население от 15 лет и старше					
Средний	121032	121170	122804	121786	116489
Высокий		121285	123339	123624	120036
Низкий		121055	122269	119947	112942
Без миграции		120643	121275	118163	110613
Рабочая сила	74745				
Средний		75441	76638	75486	68929
Высокий		75524	77044	76874	71532
Низкий		75357	76233	74099	66327
Без миграции		75067	75491	72781	64686
Занятые	74745				
Средний		730567	74046	73122	66874
Высокий		73137	74437	74465	69394
Низкий		72976	73656	71779	64353
Без миграции		72698	72941	70502	62763



(Б) ПРОГНОЗ НАСЕЛЕНИЯ МОСКВЫ ПО УРОВНЮ ОБРАЗОВАНИЯ

Специфика прогноза: «передвижка» высоких уровней образования по когортам.

Два вида оценок: (1) Образовательные контингенты (2) уровень образования

Основная гипотеза

Прогноз выполняется по когортам.

Уровень образования каждой когорты повышается за счет доли лиц с высшим образованием.

Целевые значения определялись с учетом максимально достигнутого уровня образования в когортах и современными тенденциями в развития системы образования (сдвиг в сторону среднего профессионального образования).

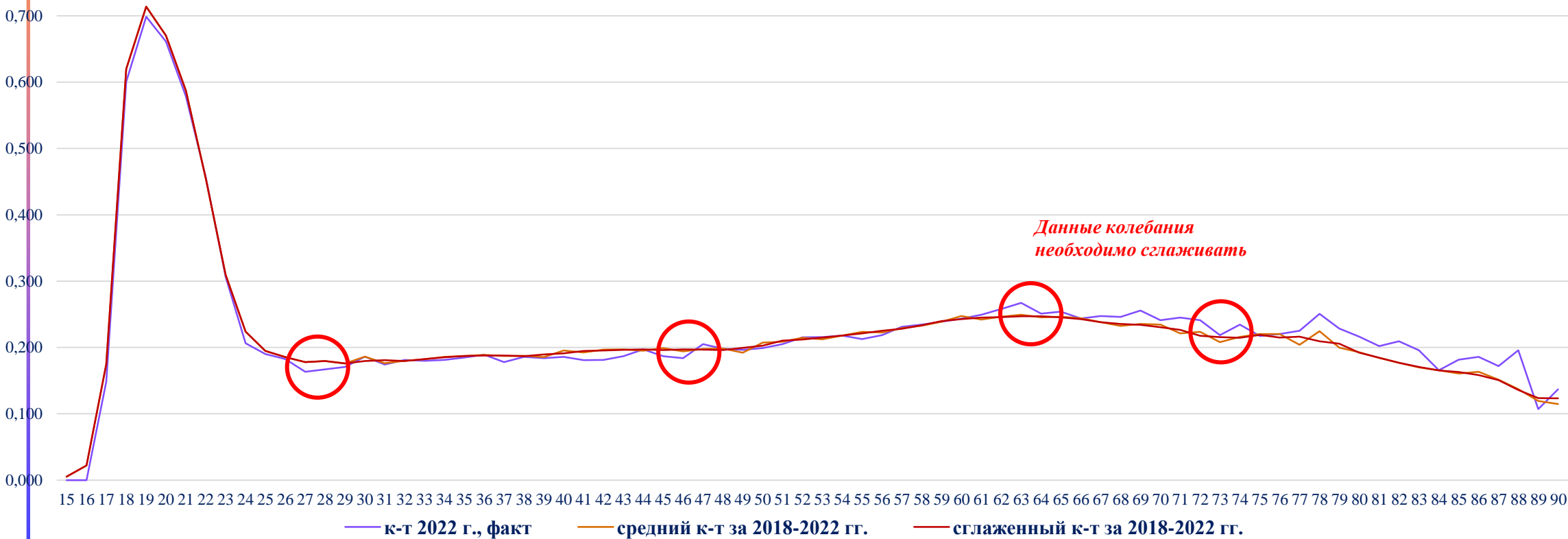
ПРОГНОЗИРОВАНИЕ ОХВАТА УРОВНЯ ОБРАЗОВАНИЯ ПО ДАННЫМ ОРС

Сглаживание данных

Проблема: Значения по однолетним интервалам подвержены колебаниям из-за объема выборки.

Предложения: Взять среднее значение охвата образования **по возрасту** за 5 лет. Сгладить значения методом скользящего среднего.

Уровень охвата средним образованием, мужчины



ПРОГНОЗИРОВАНИЕ ОХВАТА УРОВНЯ ОБРАЗОВАНИЯ ПО ДАННЫМ ОРС

Определение нижних возрастных границ для изменения каждого уровня образования

Проблема: необходимо определить четкие значения возрастов нижних границ.

Предложения: определить возрастные границы следующим образом:

Не имеют образования старше 21 года

Основное общее образование старше 23 лет

Среднее общее образование старше 25 лет

Среднее профессиональное образование старше 21 года

Высшее образование старше 22 лет

Определение верхних возрастных границ для изменения каждого уровня образования

Проблема: неизменный уровень образования в когортах старше 30 лет приводит к резкому скачку от целевого значения.

Предложения: добавить переходный период для возрастов 30-32 года (25, 50, 75% от целевого значения).

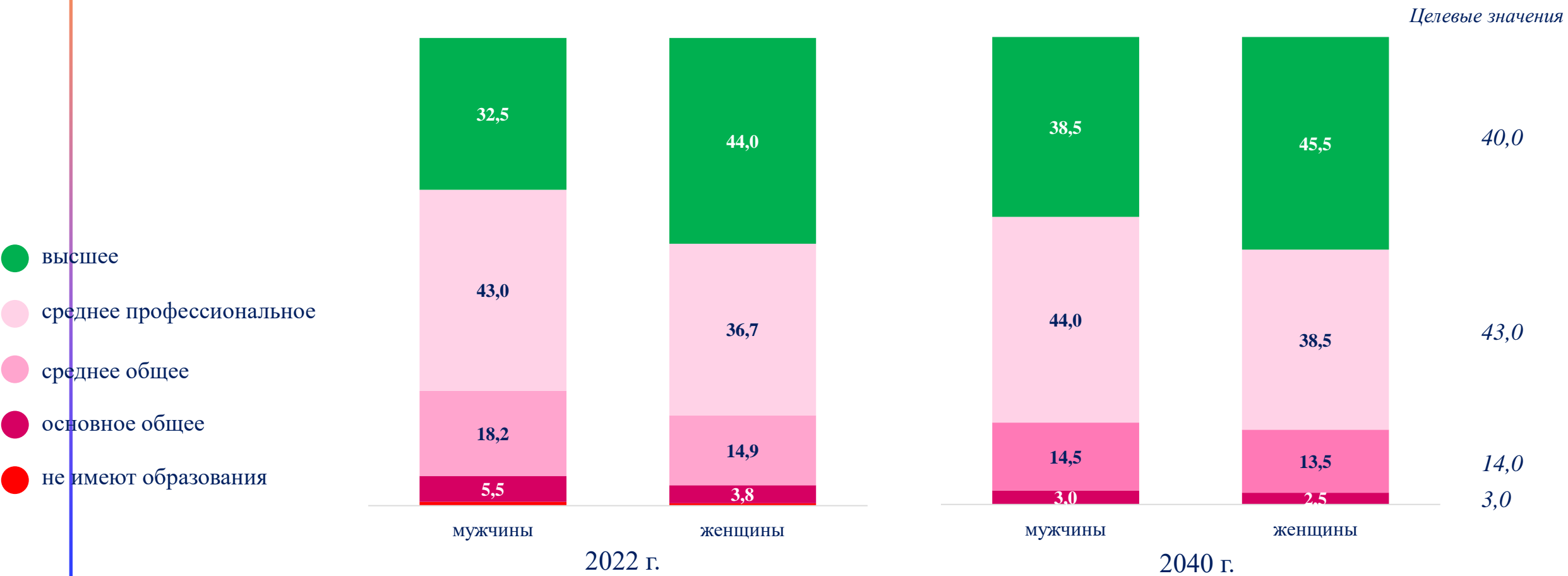
ПРОГНОЗИРОВАНИЕ ОХВАТА УРОВНЯ ОБРАЗОВАНИЯ ПО ДАННЫМ ОРС

Определение коэффициентов изменения охвата для каждого уровня образования

Проблема: необходимо определить коэффициенты изменения отдельно для мужчин и для женщин.

Предложения: определить следующие коэффициенты:

Структура по уровню образования в возрастах 25-30 лет (*группа, которая полностью подвержена изменениям*), %



ПРОГНОЗИРОВАНИЕ ОХВАТА УРОВНЯ ОБРАЗОВАНИЯ ПО ДАННЫМ ОРС

Способ прогнозирования:

Для когорт старше 30 лет – метод передвижки возрастов

Для возрастов 15-21 (23) лет охват остается неизменным

Для возрастов 21 (23) – 20 лет применяется повышающий/понижающий коэффициент с учетом передвижки возрастов



ПРОГНОЗИРОВАНИЕ ОХВАТА УРОВНЯ ОБРАЗОВАНИЯ ПО ДАННЫМ ОРС

Расчет коэффициента повышение/снижения уровня охвата:

$$K = \frac{\text{текущий охват} - \text{целевой охват}}{\text{количество прогнозных лет}}$$

ПРОГНОЗИРОВАНИЕ ОХВАТА УРОВНЯ ОБРАЗОВАНИЯ ПО ДАННЫМ ОРС

Определение коэффициентов изменения охвата для каждого уровня образования по г. Москве

Проблема: необходимо определить коэффициенты изменения отдельно для мужчин и для женщин.

Предложения: определить следующие коэффициенты:

Структура по уровню образования в возрастах 25-30 лет (группа, которая полностью подвержена изменениям), %

Уровень образования	2022 г.		2040 г.		Комментарии
	мужчины	женщины	мужчины	женщины	
не имеют образования	0,1	0,0	0,0	0,0	Снизится до 0
основное общее	0,7	0,4	0,5	0,3	Снизится, но должно быть пороговое значение. По РФ цель 3%.
среднее общее	7,2	7,3	5,5	5,6	По РФ цель 14%, в Москве в 2,5 р. ниже
среднее профессиональное	32,7	33,5	24,0	24,6	Снизится ли в пользу высшего образования? Учитывая развитие сектора услуг, то там СПО достаточно. Если растёт сектор услуг, СПО не будет снижаться? Высшее образование дорого стоит, тогда СПО будет расти?
высшее	59,3	58,8	70,0	69,5	Цель по РФ – 40%. Доля высшего образования будет увеличиваться, необходимо оценить насколько. Незначительный рост в 3-5% или существенный?



Анализ субъектов и методологические предложения:

Количество студентов высшего образования

В возрасте 17-27 лет в расчете на 1000 человек населения соответствующего возраста, 2021 г.

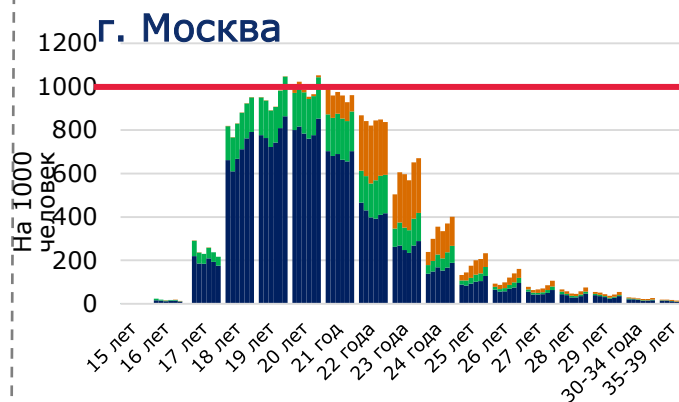
Рейтинг регионов по охвату образованием



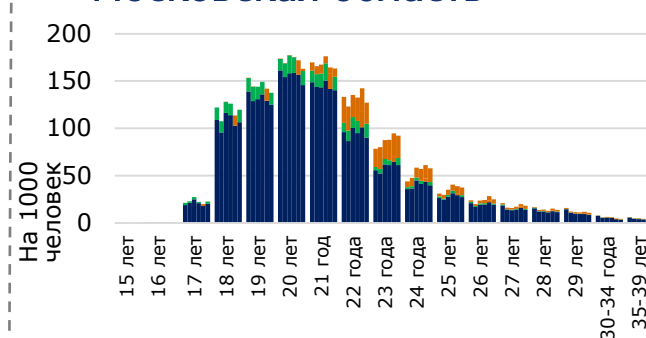
Методологические особенности расчетов

Охват высшим образованием

На 1000 человек соответствующего возраста

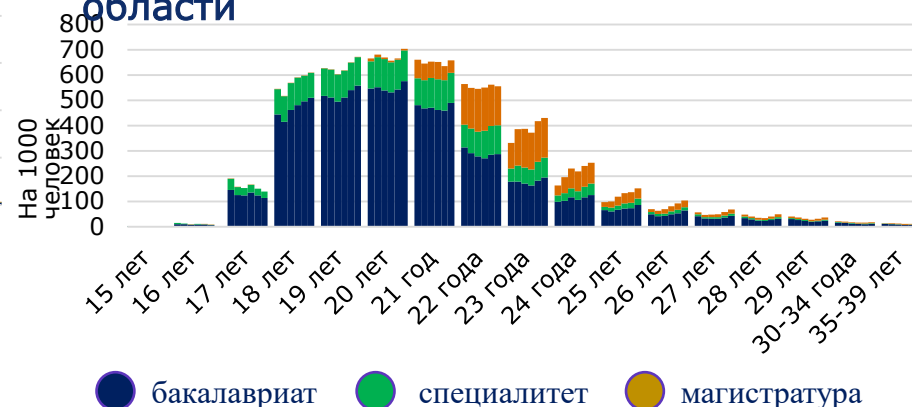


Московская область



Высокие значения по охвату в г. Москве свидетельствуют о недоучете численности населения в соответствующих возрастах и высокой маятниковой миграции, что приводит к необходимости рассмотрения значений в границах Московской агломерации

В среднем по г. Москве и Московской области



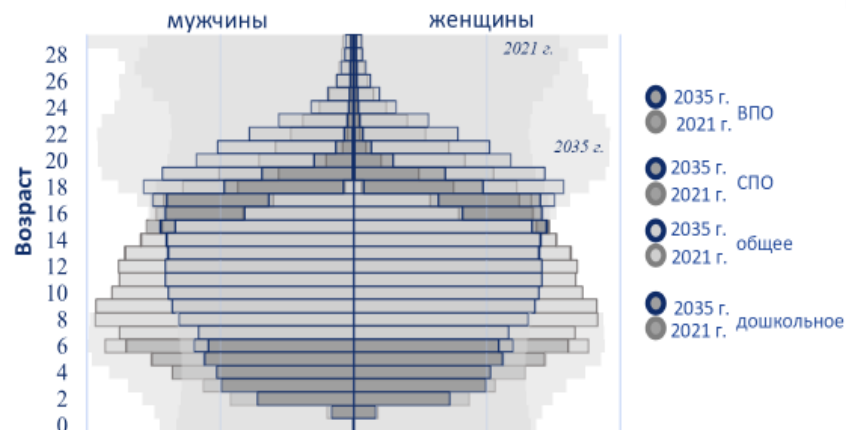


Результаты прогнозирования:

По количеству учащихся всего, человек

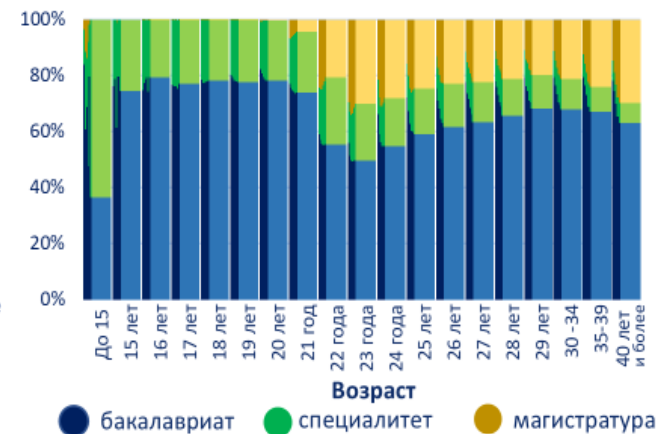


В структуре населения, человек, %



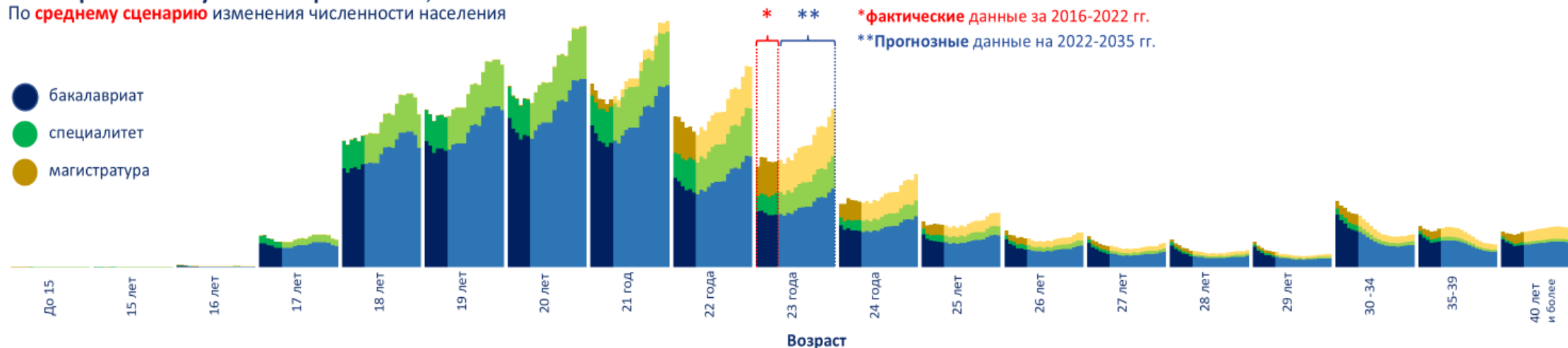
По структуре образования, %

По **среднему сценарию** изменения численности населения



По возрастам и ступеням образования, человек

По **среднему сценарию** изменения численности населения



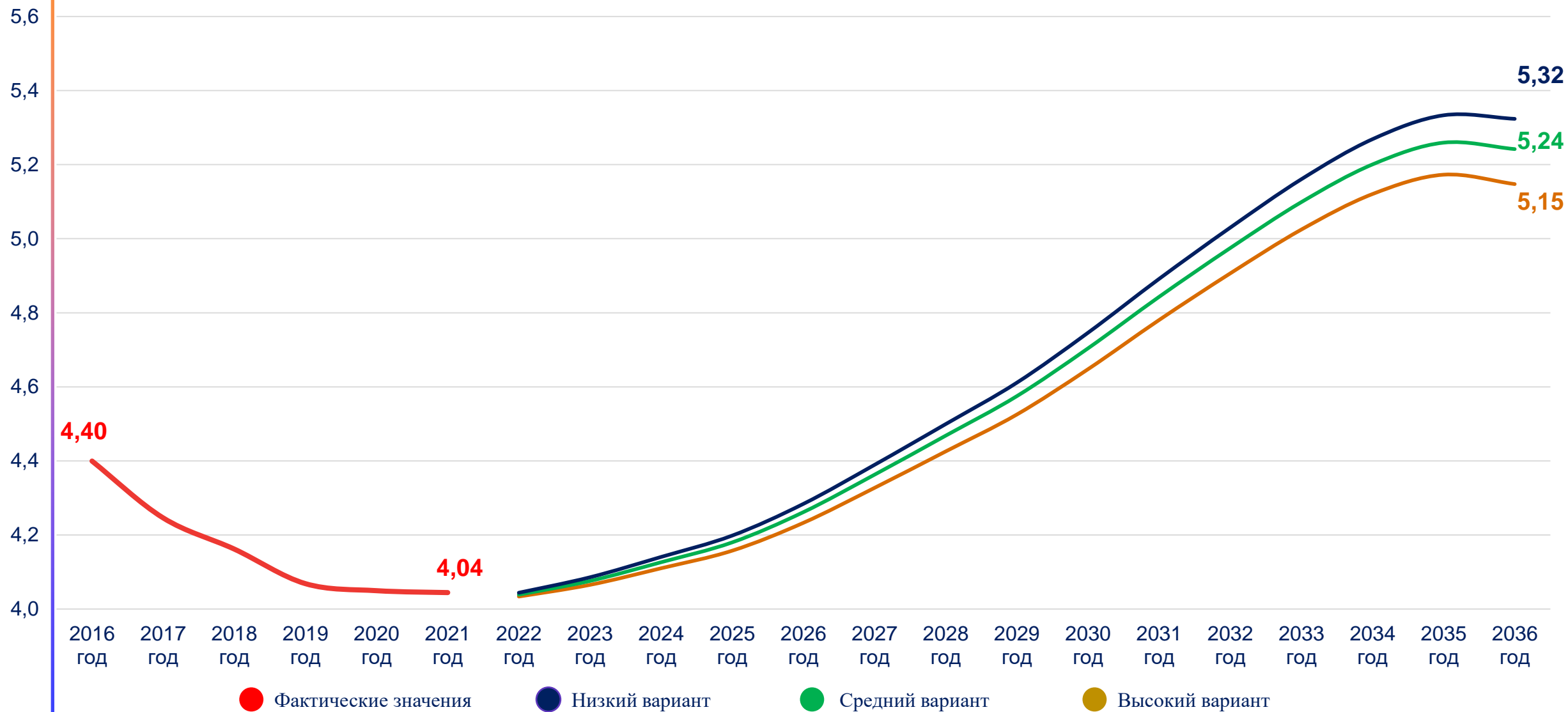


(В) ОБРАЗОВАТЕЛЬНЫЕ КОНТИНГЕНТЫ В РОССИИ

КОЛИЧЕСТВО СТУДЕНТОВ ВЫСШЕГО ОБРАЗОВАНИЯ ПО РОССИЙСКОЙ ФЕДЕРАЦИИ

2016-2021 гг. фактические значения, 2022 – 2036 гг. прогнозные значения, человек

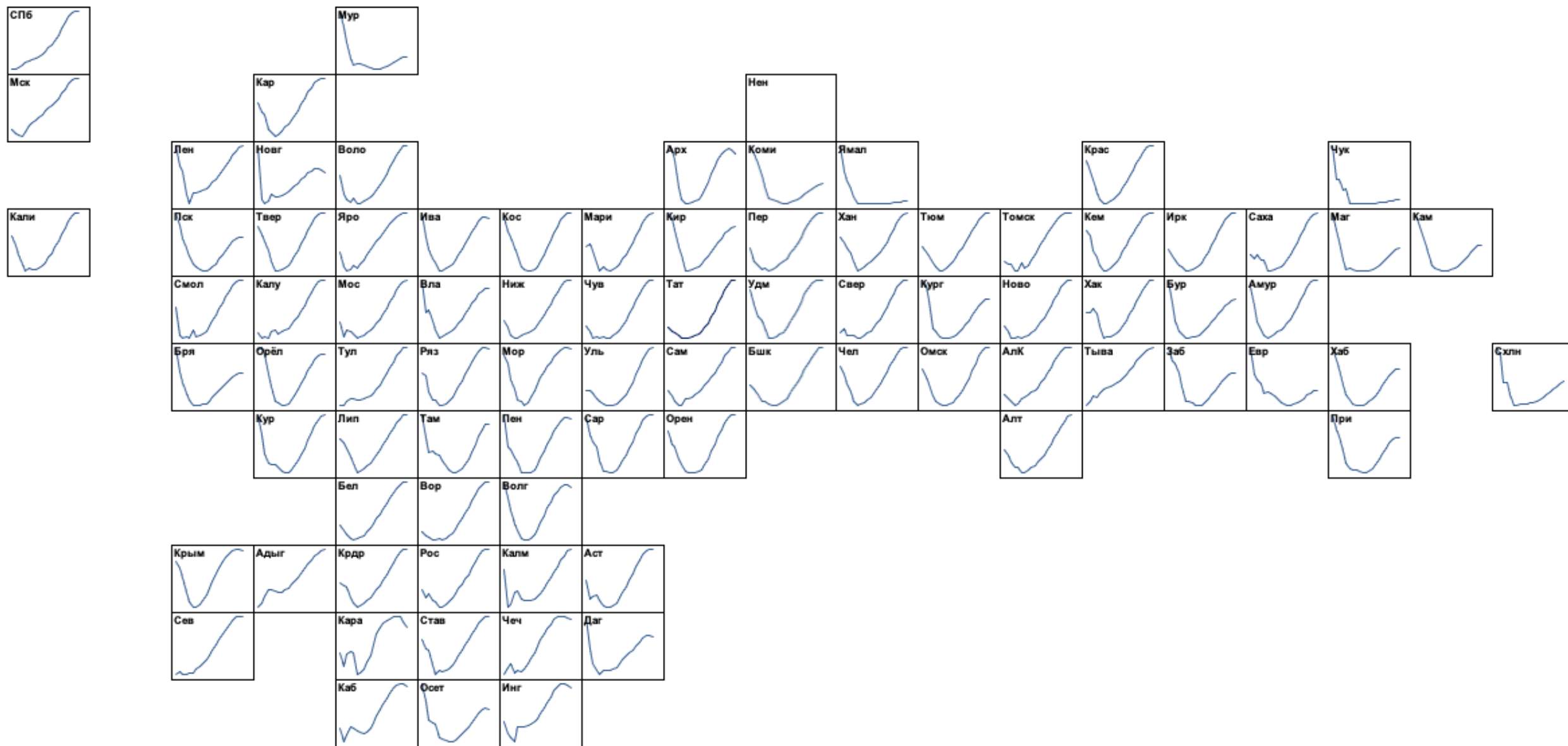
При условии, что охват останется на уровне 2021 г., млн человек



ДИНАМИКА СТУДЕНТОВ ВЫСШЕГО ОБРАЗОВАНИЯ ПО СУБЪЕКТАМ Российской Федерации

2016-2021 гг. фактические значения, 2022 – 2036 гг. **средний сценарий** прогноза, человек

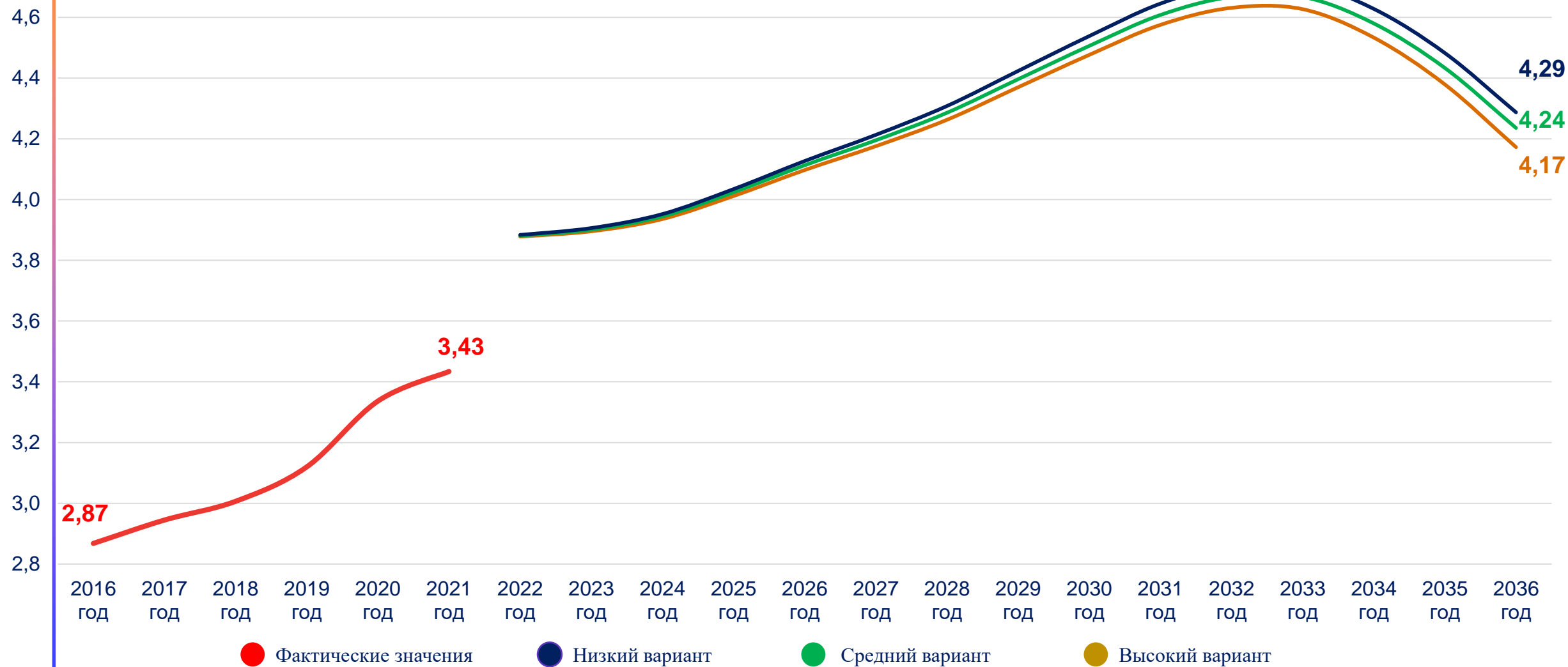
При условии, что **охват** останется на уровне 2021 г.



КОЛИЧЕСТВО СТУДЕНТОВ СРЕДНЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ В РОССИЙСКОЙ ФЕДЕРАЦИИ

2016-2021 гг. фактические значения, 2022 – 2036 гг. прогнозные значения, человек

При условии, что охват останется на уровне 2021 г., млн человек

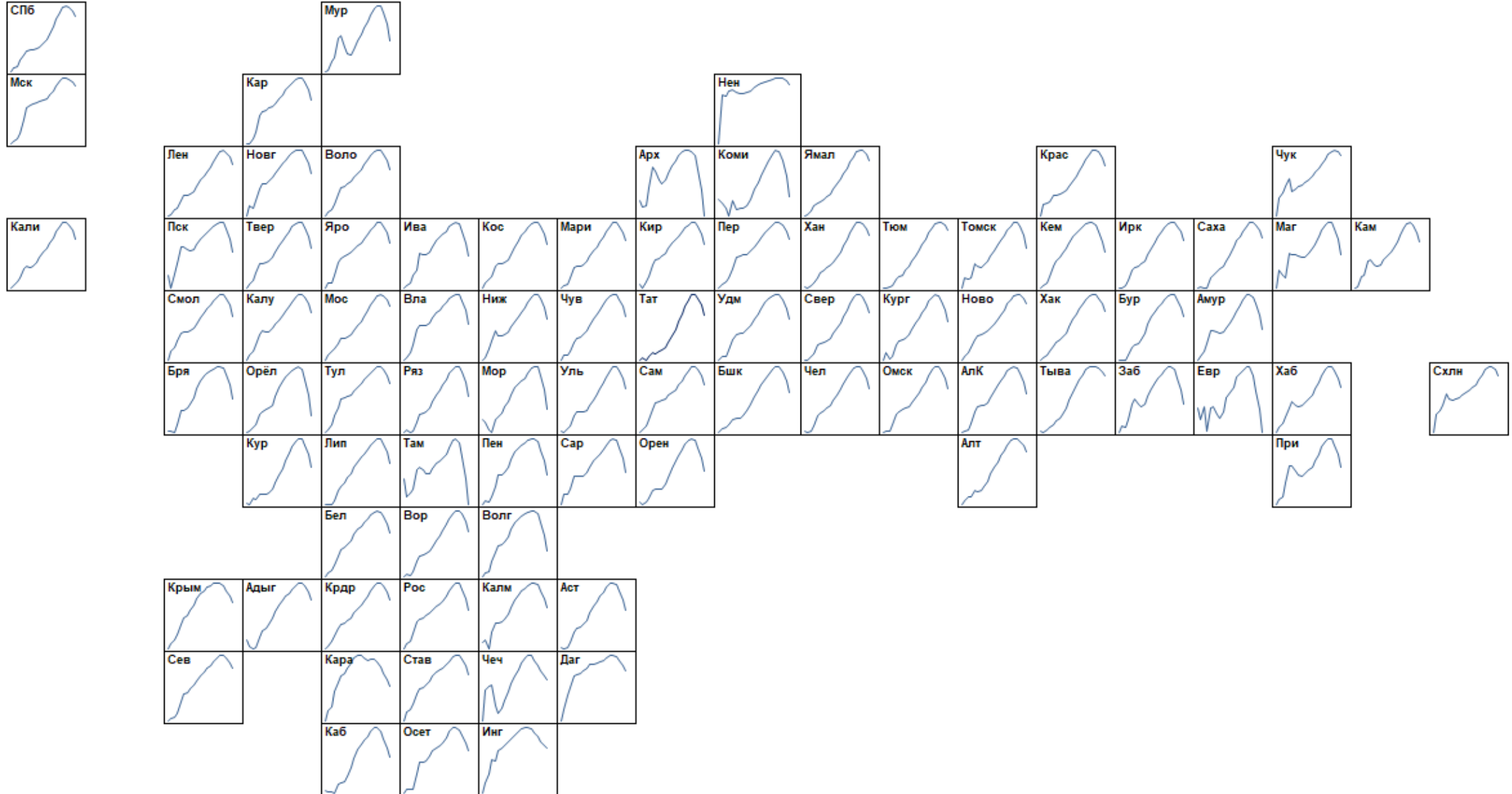


ДИНАМИКА СТУДЕНТОВ СРЕДНЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ ПО СУБЪЕКТАМ

Российской Федерации

2016-2021 гг. фактические значения, 2022 – 2036 гг. **средний вариант** прогноза, человек

При условии, что **охват** останется на уровне **2021 г.**





(Г) АГЕНТ-ОРИЕНТИРОВАННАЯ ВЫЧИСЛИТЕЛЬНАЯ ДЕМОГРАФИЯ

Новое направление, требующее большого объема информации

Содержание моделей

Агент-ориентированные вычислительные модели предполагают изучение поведения людей и проверяют, могут ли правила поведения. Действующие на микроуровне, микроуровня объяснить закономерности макроуровня.

Ключевой момент заключается в использовании вычислений как средства для развития теорий человеческого поведения. Основной акцент делается на объяснении, а не на предсказании поведения. Модели, в основе которых данные о поведении индивидуумов (агентах) называют агент-ориентированными моделями.

При их реализации имитируется поведение множества людей (супружеских пар). Такой подход стал возможен в связи с развитием вычислительной техники.

Прогнозы – дополнительный продукт

Агент-ориентированный и байесовский подходы

Billari, F.C., Prskawetz, A. (eds) Agent-Based Computational Demography. Contributions to Economics. Physica, Heidelberg. 2003

Hooten, M. , Wikle, C. , and Schwob, M. (2020) Statistical Implementations of Agent-Based Demographic Models. International Statistical Review, 88: 441–461

Billari, F.C. & Prskawetz, A. (2012). Agent-Based Computational Demography: Using Simulation to Improve Our Understanding of Demographic Behaviour. Springer

Критерий	Агент-ориентированные модели	Байесовские вероятностные модели
Подход	«Снизу вверх» (Bottom-up). Моделирует поведение (траектории) отдельных агентов.	«Сверху вниз» (Top-down). Анализирует агрегированные тенденции.
Основная цель	Понимание механизмов и проверка теорий. Объяснение, как возникают макроуровневые явления.	Прогнозирование и оценка неопределённости. Определение того, что вероятнее всего произойдёт.
Ключевой результат	Качественное понимание сложных систем, возникающих эффектов, сценарии «что, если».	Количественные вероятностные интервалы для будущих демографических показателей
Связь с данными	Часто используется для теоретических экспериментов. Может калиброваться на реалистичных данных.	Строго оценивается на основе исторических данных с чёткой статистической моделью.
Вычислительная сложность	Высокая. Зависит от числа агентов и сложности их правил поведения.	Высокая, но другого рода (MCMC-сэмплирование сложных статистических распределений).

Подходы дополняют друг друга



ПРОГНОЗИРОВАНИЕ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Искусственный интеллект несомненно поможет с выполнением рутинной и сложной работы по обработке данных, подбору модели и выполнению расчетов, но он не выйдет за пределы прежних трендов. Так или иначе, но концептуальное видение новых тенденций демографического развития пока остается за демографом-профессионалом.