

## Глава 4. Методы расчета численности и возрастно-полового состава населения

В данной главе мы постараемся описать методы расчета численности и возрастно-полового состава населения, использованные в данной работе. Во многом справедливые критические высказывания в наш адрес, прежде всего, прозвучавшие на прошедшем в Торонто 27-29 января 1995 г. семинаре "Население СССР в 20-ые - 30-ые годы", заставляют нас вновь вернуться к методам оценки населения СССР, использованным в нашем предыдущем исследовании (АДХ, 1993) и, в силу разных причин, недостаточно подробно там освещенным. С другой стороны, параллельное изложение метода расчета по территории, практически закрытой с точки зрения миграции (какой был СССР в 1927-1940 и 1946-1958 гг.), и по территории, связанной с другими регионами миграционным обменом, позволит лучше объяснить сложности расчета во втором случае.

### 4.1. Общие принципы

Основные использованные в данной работе методы расчета численности и возрастно-полового состава населения - это прямая и обратная передвижка по возрастам. В их основе лежит достаточно очевидное соотношение:

$$\begin{aligned} P_{y,t+1,s} &= P_{y,t,s} - D_{y,t,s} + U_{y,t,s} - V_{y,t,s}, \text{ при } y < t \\ P_{y,t+1,s} &= B_{t,s} - D_{y,t,s} + U_{y,t,s} - V_{y,t,s}, \text{ при } y = t, \end{aligned} \quad (1)$$

где  $P_{y,t+1,s}$ ,  $P_{y,t,s}$  - численность родившихся в году  $y$ , живущих на данной территории в начале года  $t$ ,  $t+1$ , соответственно,  $s$  - признак пола:  $s = 1$  соответствует мужчинам,  $s = 2$  - женщинам.  $B_{t,s}$  - число родившихся в году  $t$ ,  $D_{y,t,s}$  - число умерших в году  $t$  из числа родившихся в году  $y$ ,  $U_{y,t,s}$  и  $V_{y,t,s}$  - численность прибывших на данную территорию и соответственно выбывших из данной территории в году  $t$  из числа родившихся в году  $y$ . Число родившихся в году  $y$ , доживших до начала года  $t$ , равно числу лиц в возрасте  $a$  на начало года  $t$  ( $P_{a,t,s}$ ), где  $a = t - y - 1$ .

Формула (1) и есть формула прямой передвижки по возрастам. Уравнение обратной передвижки выглядит следующим образом:

$$\begin{aligned} P_{y,t,s} &= P_{y,t+1,s} + D_{y,t,s} - U_{y,t,s} + V_{y,t,s}, \text{ при } y < t \\ B_{t,s} &= P_{y,t+1,s} + D_{y,t,s} - U_{y,t,s} + V_{y,t,s}, \text{ при } y = t, \end{aligned} \quad (2)$$

Обычно старшая по возрасту когорта не является однолетней, а включает всех родившихся в году  $y$  и ранее. При прямой передвижке на каждом шаге численность когорт в возрасте 100 лет 101 год и старше складывается, и в дальнейшем вновь рассматриваются когорты, достигшая на начало года  $t+1$  возраста 100 лет и старше, то есть совокупность родившихся в году  $y+1$  и ранее. При обратной передвижке приходится проделывать противоположную процедуру. Для этого мы условно разделяем численность лиц в возрасте 99 лет и старше на начало года  $t-1$  на две когорты, пропорционально численностям лиц в возрастах 99 лет и 100 лет и старше на начало года  $t$ . Такой прием не вполне корректен, но возможная ошибка практически не влияет на результаты расчетов.

Авторы многих ретроспективных расчетов населения предпочитают ограничиваться определением общей численности населения. Уравнение демографического баланса без учета возраста выглядит следующим образом:

$$P_{t+1} = P_t + B_t - D_t + U_t - V_t, \quad (3)$$

где  $P_t$ ;  $B_t$ ;  $D_t$ ;  $U_t$ ;  $V_t$  - соответственно, общая численность населения на начало года  $t$ , число родившихся, умерших, прибывших, выбывших в течение года  $t$ . Возможность сопоставлять результаты расчетов по соседним возрастным группам - главное преимущество таких расчетов перед оценками общей численности, где внутренний контроль если не вовсе отсутствует, то значительно слабее.

Формулы (1) и (2) содержат числа событий в когортах в течение данного календарного года. Данные демографической статистики в СССР и России разрабатывались, как правило, по возрасту. Переход от чисел умерших, распределенных по возрасту, к числам умерших, распределенных по когортам по году рождения, осуществлялся по простейшей формуле, основанной на гипотезе о равенстве чисел умерших из двух когорт в одной возрастной группе:

$$D_{y,t,s} = 1/2 \times (D_{a,t,s} + D_{a+1,t,s}), \quad (4)$$

год рождения когорты равен  $y=t-a-1$ ,  $t$  - текущий год,  $a$  - возраст,  $D_{y,t,s}$  - число умерших в данном году  $t$  из когорты рождения года  $y$ , а  $D_{a,t,s}$  - в возрасте  $a$ . Эта формула используется для возрастов от 2 до 99 лет. Для возраста 0 лет мы действовали в соответствии с классической формулой коэффициента младенческой смертности Ратса: предполагалось, что  $2/3$  умерших до 1 года родились в

данном и  $1/3$  в предыдущем году. После внесения поправок на недоучет, это соотношение уточнялось исходя из формулы Кей-фица (Preston et al., 1972, p. 12), согласно которой средний возраст смерти на первом году жизни равен  $0,07 + 1,7 \times m_0$ , где  $m_0$  - коэффициент смертности в возрасте до 1 года.

Умершие в возрасте 100 лет и старше распределялись между возрастными группами 99 лет и 100 и старше лет пропорционально их численностям.

Аналогично выглядит формула перехода от чисел прибывших (выбывших), распределенных по возрасту, к этим же величинам, распределенным по году рождения.

Возрастные данные о мигрантах разрабатывались в исследуемый период по агрегированным возрастным группам (см. раздел 3.2). Числа умерших разрабатывались по одногодичным интервалам возраста, но аккумуляция в данных об умерших была столь значительной, что мы сочли целесообразным в возрастах старше 20 лет пользоваться только данными для пятилетних возрастных групп. Поэтому перед расчетом данные интерполировались для получения показателей по одногодичным возрастам. Для интерполяции использовалось уравнение параболы второй степени, коэффициенты которого определялись по числу событий в трех последовательных интервалах возраста. Число мигрантов в возрасте 60 лет и старше распределялось по однолетним возрастам пропорционально ранее определенным возрастным численностям населения на начало или на конец года.

Если в данных наблюдалась значительная аккумуляция по возрасту, то проводилось выравнивание методом 3-летней скользящей взвешенной средней арифметической с весами (1, 2, 1).

Важным понятием, используемым при описании методов расчета, является коэффициент недоучета в статистике населения того или иного демографического события. Коэффициент недоучета  $k^Z_t$  в году  $t$ , где  $Z=B$  соответствует родившимся,  $Z=D$  - умершим и т.д., рассчитывается как

$$k^Z_t = 1 - Z^o_t / Z_t, \quad (5)$$

где  $Z^o_t$  - число зарегистрированных событий, а  $Z_t$  - число произошедших событий. При работе с данными об умерших и мигрантах

используется не только общий, но и возрастной коэффициент недоучета  $k^Z_{a,t}$ .

В случае наличия недорегистрации, величина  $k^Z$  лежит между 0 (все события регистрируются) и 1 (полное отсутствие регистрации). В принципе,  $k^Z$  может быть меньше 0, что соответствует ситуации двойного (многократного) счета.

На практике мы сталкиваемся с следующей задачей: известна величина  $Z^0_t$  и мы пытаемся оценить  $k^Z_t$  и восстановить  $Z_t$ , которое рассчитывается по формуле

$$Z_t = Z^0_t / (1 - k^Z_t) \quad (6)$$

Величину  $K^Z_t = 1/(1 - k^Z_t)$  можно назвать поправочным коэффициентом. В отличие от коэффициента недоучета, величина  $K^Z_t$  всегда больше 1, если исключить случаи двойного счета.

#### 4.2. Расчеты по СССР

В период 1927-1959 гг. население СССР было практически закрыто для внешней миграции и, поэтому, мы сочли возможным в наших расчетах допустить ее полное отсутствие. Уравнение (2) обратной передвижки можно переписать следующим образом:

$$\begin{aligned} P_{y,t,s} &= P_{y,t+1,s} + D^o_{y,t,s} \times K^D_{a,t,s}, \text{ при } y < t \\ B^o_{t,s} \times K^B_t &= P_{y,t+1,s} + D^o_{y,t,s} \times K^D_{0,t,s}, \text{ при } y=t, \\ &\text{при } t_1 \leq t < t_2 \end{aligned} \quad (7)$$

Начнем с периода 1927-1936 гг. В тех случаях, когда правая и левая часть одного из уравнений системы соотношений (7) известна из независимых источников, то мы получаем уравнение, несущее определенную информацию о входящих в него переменных.

Если бы мы обладали данными о числе умерших по полу и возрасту и родившихся по полу, то задача расчета численности на начало каждого года между переписями 1937 и 1926 гг. ( $1927 \leq t < 1937$ ) сводилась бы к поиску таких поправочных коэффициентов  $K^D_{a,t}$  и  $K^B_t$ , чтобы на каждом шаге расчета выполнялось второе из уравнений (7), т.е. две оценки числа родившихся совпадали, а расчетное население на начало 1927 г. совпадало бы с аналогичными данными, полученными на основе переписи населения 1926 г. Итак мы имели бы 222 уравнений с 2040 неизвестными, чего явно

недостаточно для определения поправочных коэффициентов. С другой стороны, даже при полном учете движения населения, эти 222 соотношения не выполнялись бы абсолютно точно. Причины понятны: во-первых, демографические данные, полученные из разных источников, не могут быть абсолютно сопоставимы и, во-вторых, трудно надеяться, что операции выравнивания и коррекции, необходимость которых очевидна, выполнены абсолютно согласовано. Итак, с одной стороны, нам требуются дополнительные гипотезы, чтобы задача стала определенной и, с другой стороны, мы знаем, что заведомо истинные соотношения никогда не реализуются абсолютно точно на реальных данных. Система из  $N$  линейных уравнений с  $N$  переменными имеет решение только в рамках математики. Попытка оценить 10 параметров на основе 10 соотношений в реальной задаче (демографической, экономической или физической) скорее всего обречена на провал. Почерпнутые из реального мира коэффициенты уравнений содержат, быть может, и не очень большие случайные ошибки. Процесс решения уравнения перенесет эти ошибки на искомые параметры, возможно, увеличит их, придаст этим ошибкам роль "закона". Чтобы избежать этого, надо иметь избыток уравнений, что позволяет в рамках разного рода специальных процедур искать свободные от ошибок решения. Единственный выход - уменьшить количество переменных, причем уменьшить кардинально, так чтобы число уравнений стало много больше числа переменных.

Чтобы уменьшить число неизвестных, мы попытались сформулировать некоторые гипотезы о соотношении недорегистрации смертности в различных группах населения, подобно тому, как при проверках полноты регистрации ЦСУ СССР выделялись умершие в возрастах до 1 года и старше 1 года. Однако, для наших целей подобная группировка не подходит, так как в предвоенный период недоучет смертности в различных возрастах имел разную природу. В частности, в период репрессий значительная часть умерших в тюрьмах и лагерях не регистрировалась в общем порядке, что определяло значительный недоучет смертности взрослых мужчин. Поэтому мы допустили, что для возрастов до 14 лет уровень

недорегистрации не зависит от пола и равен уровню недоучета для женщин. Принятая группировка предусматривала использование следующих коэффициентов коррекции чисел умерших:

- в возрасте до 1 года, независимо от пола;
- мужчины в возрасте 1-14 лет и женщины старше 1 года;
- мужчины в возрасте 15 лет и старше.

Кроме того, мы допустили, что показатель недоучета родившихся, который, как известно, много ниже чем коэффициент недоучета умерших (см., например, табл. 3.4), не зависит от пола.

Благодаря принятым допущениям, система из 222 уравнений с 2040 неизвестными превратилась в систему из того же числа уравнений, но с 40 неизвестными. Расчет же между переписями 1937 и 1939 гг. в рамках этой схемы сводился к решению 206 уравнений с 8 неизвестными. Ни та, ни другая система уравнений не имеет решения, но могут быть найдены оценки коэффициентов коррекции в наибольшей степени, с точки зрения избранного критерия, удовлетворяющие системе уравнений.

Для расчета населения система соотношений (7) должна выполняться точно, поэтому полученный вариант оценки динамики чисел родившихся и умерших требовал дополнительной коррекции, которая сводилась к, так называемой, балансировке двухмерной таблицы.

Требования к методу балансировки: не должно меняться общее число умерших в данном году, число умерших в каждой когорте родившихся до 1927 г. должно равняться разности численности когорты в начале и в конце периода, для когорт родившихся в 1927-1936 гг. - разности численности родившихся и численности когорты в конце периода. Задача балансировки двухмерной таблицы - достаточно типичная статистическая задача имеет апробированные алгоритмы решения.

Процесс балансировки нарушает принятые гипотезы о соотношении недоучета. То есть априорные оценки уровня недоучета (до балансировки) не совпадают с его апостериорными оценками. Отсюда возникает задача поиска таких априорных оценок недоучета, чтобы, с одной стороны, они в наибольшей степени удовлетворяли системе уравнений, полученной из соотношений (7), а с другой, чтобы априорные оценки недоучета менее всего отличались от апостериорных. Помимо основных, приемлемое приближенное

решение системы соотношений (7) должно обладать следующими свойствами:

I. Недорегистрация смертей в относительном выражении выше, чем недорегистрация рождений; уровень недоучета смертности мужчин в 1937-1940 гг. выше, чем женщин; недоучет смертности на первом году жизни выше, чем в последующих возрастах. Из предварительного анализа следует, что полнота регистрации снижается между 1927 и 1933 г., а затем качество данных вновь растет до 1937 г. В 1937-1939 гг. можно предполагать значительный недоучет смертности взрослых мужчин за счет нерегистрации части смертей заключенных. В 1939 г. вероятно снижение полноты регистрации родившихся и умерших в СССР (но не в России) в результате плохого учета в присоединенных западных областях Украины и Белоруссии. Наконец, полнота учета несомненно росла между 1948 и 1959 гг.

II. Динамика показателей полноты учета рождений и смертей имеет сходный характер.

III. Расчетные кривые возрастной смертности для каждого календарного года не должны противоречить общеизвестным закономерностям изменения смертности с возрастом, описанным с помощью типовых таблиц смертности или каких то иных моделей, например модели У. Брасса (1977).

IV. Динамика погодных чисел родившихся должна согласовываться с данными, полученными на основе выборочного анамнестического обследования 1960 г. (Сифман, 1974).

V. Динамика основных демографических показателей должна быть достаточно интерпретируемой и плавной, кроме периодов катастроф.

Сформулированная задача, по видимому, может быть каким-то образом формализована, в том числе, условия I - V могут быть описаны некоторыми формальными неравенствами. Задача не имеет точного решения, и можно было бы пытаться построить некоторый алгоритм поиска наилучшего приближенного решения. Возможно это не будет простая итерационная схема, поскольку процесс итераций может не сходиться. Так как априорные оценки недоучета могут рассматриваться только с определенной точностью (скорее всего как целые проценты), задача сводится к перебору пусть большого, но конечного числа вариантов.

На практике мы пошли иным путем. Первоначальные оценки полноты учета делались из содержательных соображений, то есть экспертно (АДХ, 1990а, 1993), а затем уточнялись с помощью описанной процедуры и сформулированных критериев. Для этого на основе априорных оценок рассчитывались числа родившихся и умерших за каждый год рассматриваемого периода, полученные данные балансировались, рассчитывались апостериорные оценки недоучета, проверялось выполнение условий I - V. Результаты расчетов служили основой для определения новых априорных оценок недоучета. Эти оценки определялись не вполне формальным методом, хотя и лежали в интервале между его априорными и апостериорными оценками, полученными при предыдущем шаге расчета.

Для периода 1937-1939 гг. вообще потребовалось лишь несколько итераций, чтобы найти удовлетворительное решение.

Мы не обладали данными о числе умерших по полу и возрасту за каждый год между переписями 1926 г. и 1937 г., а за ряд лет имели только данные об общем числе умерших по полу, с выделением первого года жизни, и поэтому задача усложнялась. Для определения возрастного состава умерших использовалась известная модель смертности Брасса (1977). В качестве базовой кривой смертности для модели брались таблицы смертности населения СССР за 1926-1927 (только Европейская часть страны) и 1938-1939 гг. Кроме того, расчет учитывал известный факт эмиграции из Казахстана в период голода 1933 г. значительной группы населения - около 200 тысяч человек (Марианьский, 1969). Таким образом, окончательно для периода 1926-1937 гг. схема расчета выглядела следующим образом:

- 1) первоначальная оценка полноты учета;
- 2) расчет чисел родившихся и умерших, на основе априорных оценок недоучета и с использованием модели Брасса;
- 3) внесение поправок на эмиграцию;
- 4) балансировка;
- 5) расчет апостериорных оценок недоучета и других показателей;
- 6) проверка выполнения условий I - V;
- 7) определение новых априорных оценок недоучета.



Затем расчет вновь повторялся и так до тех пор пока не были выполнены все сформулированные условия.

Эксперименты показали, что система соотношений достаточно устойчива. Попытка несколько изменить полученные оценки недорегистрации вызывает заметный дисбаланс априорных и апостериорных оценок, в процессе устранения которого мы вновь приходим практически к тем же, что и раньше, оценкам недоучета чисел умерших и родившихся.

В какой степени и почему можно надеяться, что полученные таким образом демографические показатели отражают реальный процесс? Безусловно, нельзя утверждать, что данный расчет в точности отражает динамику населения. Однако, достигнутый баланс между численностями умерших по возрастам, числами родившихся и результатами трех переписей населения позволяет считать полученные результаты весьма правдоподобным отражением реальной динамики.

Расчет динамики населения в 1939 и 1940 гг. осуществлялся методом прямой передвижки. Трудность этого расчета в том, что мы не располагали данными о населении на конец периода и были вынуждены исходить только из косвенных (неформальных) методов оценки полноты учета. Мы опирались, во-первых, на общие тенденции показателей и, во-вторых, на соотношения уровней недоучета родившихся и умерших и недоучета умерших в разных возрастных группах между собой, которые мы ранее определили для периода 1927-1938 гг.

Оценка численности населения СССР за период 1946-1959 гг. осуществлялась методом обратной передвижки. Для оценки динамики населения в послевоенные годы мы располагали распределениями умерших по полу и возрасту и данными о числах родившихся по полу за 1946-1958 гг. Мы располагали данными только одной переписи, и метод баланса не мог быть использован в полном объеме.

Чтобы получить оценки полноты учета, мы разделили события этого периода на две подсовокупности: события, происшедшие среди родившихся после 1946 г. и среди родившихся до 1946 г. Очевидно, что в первой из этих совокупностей должен быть баланс рождений, смертей и численностей соответствующих возрастов по переписи 1959 г., и, следовательно, может быть применена та же

процедура, что и для периода 1927-1937 гг. Три обстоятельства позволяют распространить полученные оценки недоучета на всю совокупность умерших:

- устойчивая связь уровней смертности детей и взрослых;
- устойчивая связь между полнотой учета умерших детей и взрослых;
- устойчивая связь между полнотой учета родившихся и умерших.

Формальное описание связи между уровнями смертности детей и взрослых было построено (как и для предыдущих периодов) с применением модели смертности Брасса. В качестве базовой кривой смертности для модели брались рассчитанные на предыдущем этапе таблицы смертности населения СССР за 1938-1939 гг. и официальные таблицы смертности 1958-1959 гг. (Всесоюзная перепись населения 1959 г. СССР).

Модель, описывающая соотношение уровней недоучета числа родившихся и умерших, а также умерших детей и взрослых, была построена на основе результатов проверок полноты регистрации родившихся и умерших, проводившихся ЦСУ СССР с 1949 по 1957 гг. с использованием материалов более поздних проверок (Дмитриева, 1975).

Схема расчета, избранная для периода 1946-1959 гг., выглядела следующим образом:

- 1) первоначальная оценка полноты учета умерших во возрастах 5 лет и старше;
- 2) оценка на основе построенной модели показателей недоучета родившихся и умерших в возрастах до 1 года и 1-4 года;
- 3) расчет чисел родившихся и умерших, на основе априорных оценок недоучета и с использованием модели Брасса;
- 4) внесение незначительных поправок на иммиграцию и эмиграцию;
- 5) балансировка данных по поколениям, родившимся после 1946 г.;
- 6) расчет апостериорных оценок недоучета и других показателей;
- 7) проверка выполнения условий I - V;
- 8) определения новых априорных оценок недоучета.

Расчет повторялся так до тех пор, пока не было достигнуто равенство априорных и апостериорных оценок и не были выполнены другие сформулированные условия. Очевидно, что отсутствие переписных данных на начало периода 1946-1959 гг. делает расчет за послевоенные годы более уязвимым, чем за 1927-1939 гг. Хуже всего защищены наши оценки за 1946-1949 гг., когда качество учета умерших было достаточно низким, особенно в период голода 1947 г. Несмотря на это, нам представляется, что полученные оценки достаточно точны. К сожалению (а может быть к счастью) в период расчета мы еще не располагали результатами статистической разработки списков избирателей по выборам в Верховный Совет СССР, Верховные советы республик и местные советы в феврале 1946 г., в феврале и декабре 1947 г., в марте 1950 г. и марте 1954 г., выполненной ЦСУ СССР.

Сравнительный анализ по полу и возрастным группам данных о числе избирателей и результатов наших расчетов (особенно если учесть, что результаты разработки списков избирателей в них не использовались в наших расчетах) показывает, что они достаточно хорошо согласуются (см. гл. 7). Это позволяет считать полученные результаты достаточно адекватным отражением реальной динамики.

#### 4.3. Особенности расчета по России

Прежде всего отметим, что при расчете по России перечень условий I -V дополняется еще одним условием:

VI. Динамика полноты учета рождений и смертей в России и в СССР (исключая, возможно 1938-1939 гг.) в целом имеет сходный характер. Также сходный характер имеет динамика уровней рождаемости и смертности.

Для периода 1927-1939 гг. были использованы методы обратной и прямой передвижки по возрастам с поправкой на неполноту регистрации. Формулы для обратной передвижки имеют вид:

$$\begin{aligned}
 P_{y,t,s} &= P_{y,t+1,s} + D_{y,t,s}^o \times K_{a,t,s}^D - W_{a,t+1,s}^p, \text{ при } y < t \\
 B_{t,s}^o \times K_t^B &= P_{y,t+1,s} + D_{y,t,s}^o \times K_{0,t,s}^D, \text{ при } y = t, \\
 &\text{при } t_1 \leq t < t_2,
 \end{aligned}
 \tag{8}$$

где  $W^p_{a,t+1,s}$  - оценка чистой миграции в течение года  $t$  по состоянию на конец года<sup>1</sup>, по полу и возрасту. Аналогично, формула прямой передвижки записывается как:

$$\begin{aligned} P_{y,t+1,s} &= P_{y,t,s} - D^o_{y,t,s} \times K^D_{a,t,s} + W^p_{a,t+1,s}, \text{ при } y < t \\ P_{y,t+1,s} &= B^o_{t,s} K^B_t - D^o_{y,t,s} \times K^D_{0,t,s}, \text{ при } y = t, \\ &\text{при } t_1 \leq t < t_2, \end{aligned} \quad (9)$$

Формулы (8) и (9) предполагают, что дети родившиеся в 1927-1939 гг., в этот период не меняют место жительства. Это допущение, конечно, достаточно условно, но при малых объемах миграции оно не влияет существенно на результаты расчета. В довоенный период население России росло под действием миграции, поэтому наше допущение может несколько (весьма незначительно) завысить число родившихся.

Для оценки чистой миграции по полу и возрасту первоначально осуществлялась прямая передвижка возрастов без учета миграции и определялся кумулятивный эффект миграции в расчете на начало 1939 г. за период между переписями  $SW_{y,1939,s}$  по формуле:

$$SW_{y,t,s} = P^{\Pi}_{y,t,s} - P^{\text{бм}}_{y,t,s}, \text{ при } t = 1939, \quad (10)$$

где  $P^{\Pi}_{y,1939,s}$  - численность населения на начало 1939 г., рассчитанная по данным переписи 1939 г.,  $P^{\text{бм}}_{y,t,s}$  - численность населения на начало года  $t$ , рассчитанная без учета миграции. Одновременно с этим, на основе расчета без учета миграции определялись и коэффициенты дожития

$$S_{y,t,s} = P^{\text{бм}}_{y,t+1,s} / P^{\text{бм}}_{y,t,s}, \text{ при } 1927 \leq t < 1938, \quad (11)$$

Кумулятивный эффект миграции в расчете на начало года  $t$  определялся с помощью обратной передвижки на основе коэффициентов дожития по формуле

$$SW_{y,t,s} = SW_{y,t+1,s} / S_{y,t,s}, \text{ при } 1927 \leq t < 1938, \quad (12)$$

Как и при расчете по СССР, каждый вариант расчета определялся набором показателей недорегистрации: родившихся (независимо от пола); умерших в возрасте до 1 года (независимо от пола); мужчин в возрасте 1-14 лет и женщин старше 1 года; мужчин

---

<sup>1</sup> То есть без учета мигрантов, умерших в год переезда. Эта оговорка - существенна, так как в число мигрантов входят депортированные и репрессированные.

в возрасте 15 лет и старше. а также показателями  $w_t$  равными доле чистой миграции данного года  $t$  в общем миграционном приросте за период 1927-1938 гг. :

$$\sum_{1927 \leq t < 1938} w_t = 1 \quad (12)$$

Ряд величин  $w_t$  использовался для определения оценки чистой миграции в течение года  $t$  по состоянию на конец года

$$W_{y,t,s} = SW_{y,t,s} \times \omega_{t,s}, \text{ при } 1927 \leq t < 1938, \quad (13)$$

где  $\omega_t$  определяется из соотношений

$$\begin{aligned} \omega_{t,s} &= (w_t / SW_{t,s}) / \sum_t (w_t / SW_{t,s}) \\ SW_{t,s} &= \sum_y SW_{y,t,s} \end{aligned} \quad (14)$$

Легко убедиться, что система соотношений (13) - (14) обеспечивает, с учетом способа определения кумулятивного эффекта миграции  $SW_{y,t,s}$ , с одной стороны, выполнение соотношений (8) и (9), при заданных коэффициентах коррекции, а с другой стороны, объемы чистой миграции  $W_{t,s} = \sum_{y,s} W_{y,t,s}$ , как и предполагалось, соотносятся как величины  $w_t$

$$W_{1927,s} : W_{1928,s} : \dots : W_{1938,s} = w_{1927} : w_{1928} : \dots : w_{1938} \quad (15)$$

Мы обладали данными о смертности по возрасту только за период начиная с 1935 г., поэтому в расчетах за более ранний период, для определения распределения умерших по возрасту, была использована регрессионная модель. В основе модели лежало предположение о сходстве возрастных показателей смертности населения России и СССР. В частности, мы предположили что вероятности смерти  $Q^P_a$  для России могут быть представлены как

$$Q^P_{a,s} = Q^{СССР}_{a,s} \times (\alpha_{t,s} + \beta_{t,s} a), \quad (16)$$

где  $\alpha$  и  $\beta$  - параметры модели. Данная модель была предложена нами (Андреев, 1979. Андреев, Добровольская, 1979) и уже применялась в моделировании демографического перехода (Баркалов, 1984).

Параметры модели определялись на основе таблиц смертности населения России и СССР за 1926-1927 гг. (Европейская часть) и 1938-1939 гг., рассчитанных по данным статистики, без поправок на недоучет. Затем они линейно интерполировались на период с 1927 г. по 1940 г., после чего мы применяли их к уже скорректированным таблицам смертности населения СССР (АДХ, 1993).

Построенная модель смертности позволила получить первую, самую грубую оценку смертности населения России для всего предвоенного периода. Но в окончательных расчетах она использовалась только для периода 1927-1932 и 1934 гг.

Для периода голода 1933 г. модель возрастной смертности была получена также на основе кривых смертности по СССР. Поскольку по нашим оценкам уровень сверхсмертности в результате голода в России был ниже чем в среднем по СССР, базовая кривая для модели была рассчитана как средневзвешенная кривых за 1932, 1933 и 1934 гг. с весами  $v_i$  равными соответственно  $(1-\delta)/2$ ,  $\delta$ ,  $(1-\delta)/2$ :

$$Q^{P_{a,s}} = (v_1 \times Q^{СССР}_{1932,a,s} + v_2 \times Q^{СССР}_{1933,a,s} + v_3 \times Q^{СССР}_{1934,a,s}) \times (\alpha_{t,s} + \beta_{t,s} \times a). \quad (17)$$

Параметр  $\delta$  - соотношение сверхсмертности в России и СССР, определялся исходя из оценки общего числа умерших в России в период голода 1933 г. (см. гл. 5).

Первоначальные оценки полноты учета и доли чистой миграции данного года в общем миграционном приросте за период определялись из содержательных соображений (гл. 5), а затем уточнялись с помощью серии расчетов. При этом мы стремились к тому, чтобы расчетная численность населения на начало 1937 г. совпадала с численностью, независимо определенной на основе переписи 1937 г.; априорные и апостериорные оценки недоучета совпадали; объемы миграции лежали в границах, определенных из качественных соображений; возрастная структура накопленной чистой миграции к моменту переписи 1939 г. не противоречила известным представлениям о возрастной дифференциации миграционной подвижности; и выполнялись сформулированные выше критерии I - VI.

При определении коэффициентов недоучета демографических событий в России мы опирались на соотношения, полученные при расчетах по СССР;

$$K^{B_t} : K^{D_{0,t}} : K^{D_{t,1}} : K^{D_{t,2}} = 1 : 1,33 : 1,55 : 1,25, \quad (18)$$

где  $K^{B_t}$  - коэффициент недоучета рождений;  $K^{D_{0,t}}$  - коэффициент недоучета младенческих смертей;  $K^{D_{t,1}}$  - коэффициенту недоучета смертности мужчин;  $K^{D_{t,2}}$  - коэффициенту недоучета смертности женщин.

Результаты каждого расчета служили основой для определения новых априорных оценок недоучета. Таким образом единственное, но весьма существенное отличие расчета по России от расчета по СССР - невозможность использования процедуры балансировки.

Для периода после переписи 1939 г. и до начала 1941 г. был использован метод прямой передвижки (9). Параметры модели определялись с учетом общих тенденций показателей и соотношений уровней недоучета родившихся и умерших и недоучета умерших в разных возрастных группах.

При расчете численности населения России в 1946-1958 гг. был использован метод обратной передвижки по возрастам в следующей модификации:

$$P_{y,t,s} = P_{y,t+1,s} + D_{y,t,ss}^o K_{a,t}^D - U_{y,t,s}^p + V_{y,t,s}^p - W_{a,t+1,s}^c, \quad \text{при } y < t \quad (19)$$

$$B_{t,s}^o \times K_{t,s}^B = P_{y,t+1,s} + D_{y,t,s}^o \times K_{0,t}^D - U_{y,t,s}^p + V_{y,t,s}^p, \quad \text{при } y = t,$$

$$t_1 \leq t < t_2,$$

где  $U_{y,t,s}^p$  - расчетное число прибывших в Россию в течение года  $t$  по полу и возрасту,  $V_{y,t,s}^p$  - расчетное число выбывших из России,  $W_{a,t+1,s}^c$  - оценка чистой миграции специальных контингентов (заключенные, ссыльные) в течение года  $t$  по состоянию на конец года.

Для периода 1946-1958 гг., когда мы располагали значительно большей информацией, использовались три поправки на недоучет умерших каждого пола: для детей до 1 года, от 1 до 4 лет и для 5 лет и старше. Отметим, что в этот период, по нашим оценкам, различия в недорегистрации умерших мужчин и женщин сократились до минимума, так как смерти заключенных начали достаточно регулярно регистрировать. В расчете использовались следующие соотношения коэффициентов недоучета умерших в разных возрастах и коэффициента недоучета родившихся, определенные на основе результатов расчета по СССР:

- для возраста 0 лет

$$K^{D_{0,t}} = 1,83 \times K^{B_t} + 0,23;$$

- для возраста 1-4 года

$$K^{D_{1-4,t}} = 1,69 \times K^{B_t} + 0,24; \quad (20)$$

- для возраста 5 лет и старше

$$K^{D_{5+,t}} = 0,69 \times K^{B_t} + 0,03,$$

где  $K^{B_t}$  коэффициент недоучета рождений.

Схема дальнейшего расчета мало отличается от аналогичной по СССР. Она учитывала указанные связи между коэффициентами недоучета, требование равенства априорных и апостериорных оценок недоучета, условия I - VI, а также согласованность результатов расчетов по возрастам с данными о числе избирателей (см. гл. 7).

Наличие интенсивной миграции между Россией и другими союзными республиками СССР делает невозможным жесткий баланс численности населения и вносит необходимость оценки миграционного прироста. Поэтому, расчет по России может быть более уязвимым для критики, чем расчет по СССР в целом. В то же время, данные о движении населения России за весь период 1927-1958 гг. значительно полнее и качественнее, чем аналогичные данные по СССР в целом.