Chapter II

INDIRECT MEASURES OF NET INTERNAL MIGRATION

Regardless of whether direct questions on migration have been asked in the census, it is possible to estimate net intercensal migration on the basis of census counts of the population of component areas at two successive censuses along with some additional information that is normally available from the censuses or from other sources.

The population increment between any two dates for any given geographic area is the result of natural increase (births minus deaths) and net migratory movement. If the country is a closed one as far as population growth is concerned, i.e., if there has been virtually no migration between the given country and other countries, then the net migratory movement for a given geographic area must be the result of internal migration, i.e., in-migration minus out-migration. Where the population is not closed, problems arise in measuring the effects of internal migration. These are dealt with in the discussion of specific techniques.

Given the population of an area at two points in time and an estimate of natural increase during the interval, we can calculate the number that would be expected at the end of the interval in the absence of migration. The difference between the observed and expected numbers at the end of the interval, or the difference between the observed and the expected change, gives an estimate of net change due to migration.

Approaches to estimating the expected population or the expected change are of two types: (*a*) through vital statistics and (*b*) through the use of estimates of the probability of survival. Applications of these approaches are discussed below.

VITAL STATISTICS METHOD (VS)

Where reliable statistics of births and deaths to the residents of each component area of a country are available, it is possible to estimate the natural increase between two census dates or between any two dates for which the population is known. The estimate of net migration is then obtained by subtracting the natural increase from the total population change. This "balancing equation" can be put in the following simple form:

$$\text{Net } M = (p_{t+n}) - p_t - (B - D) \qquad (4)$$

where for any given area Net M = net migration, p_t is the population at the earlier census, p_{t+n} is the population at the later census, B is the number of births that occurred to residents of the area during the intercensal period, and D is the number of deaths that occurred to residents of the area during the same period. An application of the formula is given in table 23.

TABLE 23. ESTIMATES OF NET MIGRATION TO MADRAS CITY BY THE VITAL STATISTICS METHOD, 1951-1961

1. Population of Madras, 1951 = p_t =	1,416,056
2. Population of Madras, 1961 = p_{t+1} =	1,729,141
3. Increase in population, 1951-1961 = (2) − (1) =	313,085
4. Number of births in Madras, 1951-1961 = B =	653,190
5. Number of deaths in Madras, 1951-1961 = D =	371,286
6. Natural increase in Madras, 1951-1961 = (4) − (5) =	281,904
7. Net migration to Madras, 1951-1961 = (3) − (6) =	31,181

SOURCE: The population figures are taken from *Census of India*, vol. IX, *Madras*, part II-A. The figures of births and deaths are taken from *Vital Statistics of India, 1962*, issued by the Registrar General, India.

This method can, of course, be applied not only to the total population of the area but to particular segments of the population with characteristics which do not change (sex): or for which the change over time is determinable (age).

Thus, if two censuses are taken n years apart, the population aged x years at the earlier census and surviving to the later census will be $x+n$ years old. If the number of deaths in the intercensal period to persons who were x years of age at the time of the earlier census can be determined, then the net migration of persons in this age cohort can be obtained. The balancing equation can then be written in the form:

$$\text{Net } M(x) = p_{x+n,t+n} - p_{x,t} + D(x) \qquad (5)$$

where $M(x)$ is the migration among persons who were aged x at the earlier census, $p_{x+n,t+n}$ is the population aged $x+n$ years at the later census, $p_{x,t}$ is the population aged x at the earlier census, and $D(x)$ is deaths among persons who were aged x at the earlier census.

The problem is that vital statistics (data on births and deaths) are unlikely to be available in the kind of detail required for the cohort approach. Deaths are usually tabulated by age at death rather than by age at a fixed date, for example, the date of the last census. Acceptable estimates of cohort deaths can be made only if the death statistics are tabulated in considerable detail, both by age and by time of death for each geographic area. Few countries are likely to have their data in such detail. Even if they do, there will be problems of incomplete coverage and of misreporting of age in both the census and the death statistics.

From equation (4), it is readily seen that there are several ways in which errors can enter into the net migration estimate. Any uncompensated errors in the enumerated population will be reflected in the estimate. Similarly, errors in vital statistics will affect the accuracy of the estimate. However, it is important to note that some of the error will cancel out in the estimating equation.

Under-enumeration of the total population is probably universal. If the amounts of under-enumeration in the two censuses are equal, there will be no error from this source in the estimate of net migration because p_{t+n} and p_t have opposite signs in the equation. Even if the errors in p_{t+n} and p_t are not equal, the amount of error in Net M is likely to be less than that in either p_{t+n} or p_t. In most countries, relative under-enumeration probably decreases from one census to the next; but since the total population usually increases between censuses, the absolute amount of under-enumeration may increase, decrease or remain more or less the same.

In the developing countries, vital statistics are generally of poor quality, and their errors will be reflected in the migration estimate. But again, a part of the error in the count of births will be cancelled by the error in the count of deaths since B and D have opposite signs in the equation. If the relative undercount is greater for deaths than for births (as in the case in many developing countries) and if the number of deaths is smaller than the number of births (as in almost always the case), there will be a tendency for the error in births to approach the error in deaths. A last point to be noted is that even if the total change $(p_{t+n}-p_t)$ and the natural increase $(B-D)$ are in error, there is some possibility that the migration estimate will be reasonably accurate. The error in natural increase will tend to cancel with the error in total change if the net census error and the natural increase error are both positive or both negative.

The above discussion has assumed that international migration is either nil or negligible. Where this is not so, the estimate of net migration derived by equation (4) will measure the combined effect of both internal and external migration rather than of internal migration only. Thus, if there is appreciable net immigration into a country and if this net immigration tends to concentrate in a particular area, then the net migration for that area as derived from the balancing equation may be principally the result of net immigration; net internal migration may even be negative. One way of partially disentangling internal from external migration would be to use the balancing equation on the native-born population, provided, of course, the natives have not experienced external migration. This procedure would require that not only the census statistics but also the death statistics be available separately for the native and the foreign-born of each area. Since it is unlikely that the death statistics would be classified by nativity, an approximate method might assume that the death rates among natives and non-natives are the same. It would then be possible to start with the native-born population at the earlier census and to deduct from it that portion of the total deaths in the area which can be attributed to the native-born population only. The balancing equation would then include, in addition to these terms, births in the intercensal period and the native population at the time of the second census. This procedure will yield only partial estimates of net internal migration, however, for the internal migration of the foreign-born who were already in the country at the beginning of the interval will be excluded.

If equation (5) is to be used to estimate net migration by age, it should be kept in mind that under-enumeration is likely to vary systemically by age and that biases are thereby introduced into the estimates.

Survival Ratio methods

The second general approach to the estimation of net internal migration for the period between two censuses involves the use of survivorship probabilities. The basic information required is the number of persons classified by age and sex as enumerated in each area at two successive censuses and a set of survival ratios which can be applied to the population at the first census in order to derive an estimate of the number of persons expected to survive to the second census. The difference between the enumerated population at the second census and the expected population is the estimate of net migration. As with the VS method, this approach provides a good measure of net internal migration only when it is applied to a population for which external migration is negligible. The procedure may be expressed symbolically as:

$$\text{Net } M'(x) = p_{x+n,t+n} - S.p_{x,t} \qquad (6)$$

where $M'(x)$ is the net migration of survivors among persons aged x at the first census in a given area (they will be aged $x+n$ at the second census), $P_{x,t}$ is the population aged x in that area at the first census, $p_{x+n,t+n}$ is the population aged $x+n$ years in the same area at the second census separated from the first census by n years, and S is the survival ratio. It yields an estimate of net change due to the migration of persons who survived to the second census.

An alternative to estimating the expected number of persons at the second census by thus applying "forward survival ratios" to the numbers in the first census would be to do the reverse, i.e., to estimate the number of persons that would have been x years of age at the earlier census from the number who are enumerated as $x+n$ years old in the second census by applying "reverse survival ratios" (the reciprocals of forward survival ratios). The rationale here is that the number of persons x years old at the earlier census is equal to the number of persons at the second census who are n years older *plus* the deaths to this cohort. The resulting estimate of net migration thus includes deaths to the migrant cohorts and is equivalent to an assumption that all migration occurred at the beginning of the interval. Symbolically:

$$\text{Net } M''(x) = \frac{1}{S} \cdot p_{x+n,t+n} - p_{x,t} \qquad (7)$$

These two procedures always give different estimates of net migration; but if one of them is known, the other can be calculated. They are related by the equation:

$$\text{Net } M'' = \frac{1}{S} \cdot \text{Net } M' \qquad (8)$$

Since S is positive and usually less than unity, Net M'' always has the same sign as Net M' but is usually greater than Net M'.

The assumption about the timing of migration that is implicit in equation (7) is not realistic. A third procedure, the Average Survival Ratio method, yields an average of Net M' and Net M''. That is:

$$\text{Net } \bar{M} = \frac{(\text{Net } M' + \text{Net } M'')}{2} \tag{9}$$

This estimate, like Net M'', includes migrant deaths but on the assumption that deaths and migrations were evenly distributed over the decade or that all migration occurred at the middle of the interval.[1] It is seen that:

$$\text{Net } \bar{M} = \frac{1+S}{2S} \cdot \text{Net } M' \tag{10}$$

Life Table Survival Ratios (LTSR)

If a life table describing the average mortality conditions of the intercensal period is available for the particular area, survival ratios may be calculated from it and used to estimate net migration for the area. The procedure is evident from the formula given below and the illustrative example given in table 24. If the intercensal period is ten years and the population is classified by five-year age groups, forward survival ratios are given by the formula:

$$_{10}S_x = {_5L_{x+10}}/{_5L_x} \tag{11}$$

where $x = 0, 5, 10\ldots$, $_{10}S_x$ is the 10-year survival ratio from age group (x to $x+4$) to age group ($x+10$ to $x+14$) and $_5L_{x+10}$ and $_5L_x$ are the numbers of persons in the age groups $x+10$ to $x+14$ and x to $x+4$ respectively in the stationary population of the life table. For the terminal age group (say 70+), the survival ratio is given by the formula:

$$_{10}S_{70+} = \frac{T_{80}}{T_{70}} \tag{12}$$

It is important to note that the life table used should measure the average mortality conditions of the intercensal period and should be reasonably applicable to the area for which migration estimates are required. If a life table is not available for the area, but the average mortality level of the period is approximately known, model life tables can be used to calculate the survival ratios.[2]

If an appropriate life table is available and if the census age data are free from error, the LTSR method should give fairly accurate estimates of net migration for persons who were still alive at the time of the second census. But usually the age data are defective, and these defects will be present in the migration estimates. Incompatibility between life table survival ratios and census age data will show itself in an irregular pattern of migration estimates by age and in the failure of the sum of net migration balances for all areal units to add to zero,

TABLE 24. ILLUSTRATION OF PROCEDURES FOR ESTIMATING NET INTERCENSAL MIGRATION BY AGE ACCORDING TO THE FORWARD LIFE TABLE SURVIVAL RATIO METHOD, MALE POPULATION OF GREATER BOMBAY, 1941-1951

Census of 1941		*Ten-year life table survival ratio*	*Census of 1951*		*Expected survivors, 1951*	*Net migration 1941-1951*
Age (1)	*Population* (2)	(3)	*Age* (4)	*Population* (5)	(6) = (2)×(3)	(7) = (5)−(6)
0-4	77,135	.9087	10-14	132,870	70,093	+62,777
5-9	85,134	.9573	15-19	170,227	81,786	+88,441
10-14	79,185	.9471	20-24	263,971	74,996	+188,975
15-19	82,603	.9308	25-29	253,964	76,887	+177,077
20-24	126,247	.9223	30-34	195,373	116,438	+78,935
25-29	155,344	.9161	35-39	151,259	142,311	+8,948
30-34	138,843	.9047	40-44	118,383	125,611	−7,228
35-39	109,356	.8850	45-49	76,421	96,780	−20,359
40-44	81,626	.8548	50-54	65,897	69,774	−3,877
45-49	47,062	.8122	55-59	32,265	38,224	−5,959
50-54	36,908	.7535	60-64	22,248	27,810	−5,562
55-59	15,134	.6726	65-69	9,655	10,179	−524
60+..........	25,094	.3866	70+	10,100	9,701	+399
All ages	1,059,911	· Total 10+		1,502,633	940,590	+562,043

SOURCE: The age data from the 1941 and 1951 censuses of India have been adjusted for area changes. The survival ratios in column (3) were derived from the United Nations model life table corresponding to a life expectancy at birth of 45 years for males. See *Manual III: Methods for Population Projections by Sex and Age* (United Nations publication, Sales No.: 56.XIII.3).

[1] For a more detailed discussion, see Jacob S. Siegel and C. Horace Hamilton, "Some considerations in the use of the residual method of estimating net migration", *Journal of the American Statistical Association* (Washington, D.C.), vol. 45, September 1952, pp. 475-500.

[2] See *Manual III: Methods for Population Projections by Sex and Age* (United Nations publication, Sales No.: 56.XIII.3), pp. 72-81, or those prepared by Ansley J. Coale and Paul Demeny, *Regional Model Life Tables and Stable Populations* (Princeton, Princeton University Press, 1966).

which it must do in each age group. Life table survival ratios are smooth, and when a set of smooth survival ratios is applied to a distorted or irregular age distribution, the resulting expected populations and net migration estimates are also distorted; and the sum of net balances for gaining areas probably will not be equal to the sum of net balances for losing areas. The discrepancy may be eliminated by forcing the sum of the positive balances to equal the sum of the negative balances, increasing one and decreasing the other, and prorating the adjustment. The lack of smoothness in the migration estimates by age may be overcome by smoothing either the census age data before applying the estimating formula or the migration estimates themselves. The former procedure is preferable, but there is one precaution that should be taken. The graduation formula should preferably be limited to narrow ranges of age; otherwise, the graduation process may distort estimates of net migration.

If a life table covering the entire intercensal period is not available, estimates of intercensal survival ratios may be obtained by averaging the survival ratios from two life tables, one applicable to the beginning and the other to the end of the intercensal period.

Sometimes it happens that the censuses are not taken at five-year or ten-year intervals, while the age data are tabulated in the usual five-year age groups. In this case, there is need to reconstruct the age distribution of one of the censuses, usually the second census, and to calculate n-year survival ratios for five-year age groups where n is the length of the intercensal period. Calculation of n-year survival ratios from a complete life table is straightforward; but if only an abridged life table is available, some method of interpolation is required.[3] The same method may be used to reconstruct the census age data as well as the survival ratios.

[3] Sprague's multipliers are useful for this purpose. (See *Manual III: Methods for Population Projections by Sex and Age*, table 51, p. 68.)

Census Survival Ratios (CSR)

Where appropriate life tables are lacking, or where use of life table survival ratios is contra-indicated for other reasons, survival ratios can be computed from census age distributions and used instead of life table ratios to estimate the expected population. A census survival ratio is simply the ratio of the population aged $x+n$ at a given census to the population aged x at the census n years earlier. Computed for a nation as a whole, for a "closed" population, the ratio is then multiplied by the population aged x in each component area at the first census; and the expected survivors are subtracted from the corresponding population enumerated at the second census to yield estimates of net migration.

Symbolically, if $p_{i,x,t}$ refers to the population in the i^{th} state in a particular age group x at the first census (time t), $p_{i,x+n,t+n}$ the corresponding population n years older (i.e., $x+n$) at the next census, (time $t+n$) and $P_{x,t}$ and $P_{x+n,t+n}$ refer to the corresponding population of the country as a whole, the estimate of net migration is given by the formula:

$$\text{Net } M_i(x) = p_{i,x+n,t+n} - \frac{P_{x+n,t+n}}{P_{x,t}} \cdot p_{i,x,t} \qquad (13)$$

$$\text{Where } P_x = \sum_i p_{i,x} \text{ for all } x$$

Population data are usually compiled by five-year age groups and the intercensal interval is usually five or ten years. In this situation, no adjustment of the basic age

TABLE 25.A. MALE POPULATION OF KOREA, BY AGE AND HYPOTHETICAL REGIONS, 1930

	Hypothetical "regions"					
Age	*A*	*B*	*C*	*D*	*E*	*Total*
0-4	387,341	345,629	242,593	364,292	321,385	1,661,240
5-9	319,916	280,509	193,967	298,044	269,189	1,361,625
10-14	266,844	238,587	166,124	257,720	224,333	1,153,608
15-19	233,194	219,419	157,457	231,313	216,816	1,058,199
20-24	183,169	181,465	131,730	193,900	170,309	860,573
25-29	145,458	151,066	103,642	159,111	132,877	692,154
30-34	158,072	146,182	104,668	159,828	137,976	706,726
35-39	138,058	128,433	89,348	139,821	123,021	618,681
40-44	127,043	112,876	79,232	117,056	112,034	548,241
45-49	100,143	98,080	71,073	101,291	90,344	460,921
50-54	83,534	79,085	59,826	83,579	73,589	379,603
55-59	68,876	63,371	46,299	65,290	61,758	305,594
60-64	51,951	46,773	37,359	50,041	49,014	235,138
65-69	41,245	33,501	26,025	36,029	34,753	171,553
70-74	25,136	22,482	18,990	25,079	22,832	114,519
75-79	10,873	10,033	8,364	11,080	9,841	50,191
80+..........	4,300	4,429	3,243	4,526	3,825	20,323
All ages........	2,345,143	2,161,920	1,539,940	2,297,990	2,053,896	10,398,889

SOURCE: Computed by grouping the data for thirteen provinces into five hypothetical regions. Data for each of the thirteen provinces and for the total country are taken from: Yun Kim, "*The Population of Korea, 1910-1945*" (unpublished Ph.D. dissertation, Australian National University, 1966), tables A1.1 and A2.2, pp. 256-362.

TABLE 25.B. MALE POPULATION OF KOREA, BY AGE AND HYPOTHETICAL REGIONS, 1935

Age	Hypothetical "regions" A	B	C	D	E	Total
4-0	416,072	397,188	273,698	410,782	366,387	1,864,127
5-9	346,195	307,585	209,836	323,481	290,967	1,478,064
10-14	300,455	269,740	186,556	284,260	260,799	1,301,810
15-19	238,336	226,633	157,428	241,441	216,476	1,080,314
20-24	197,798	207,210	145,977	215,981	192,782	959,748
25-29	167,854	174,724	123,879	183,108	161,980	811,545
30-34	136,517	143,037	96,457	147,952	128,605	652,568
35-39	149,797	138,439	97,931	149,768	133,372	669,307
40-44	128,711	118,207	82,142	127,818	115,733	572,611
45-49	118,577	103,943	72,930	107,301	105,963	508,714
50-54	88,760	86,605	63,208	88,696	81,851	409,120
55-59	73,401	69,461	52,903	72,795	65,485	334,045
60-64	55,688	51,773	37,870	52,576	50,298	248,205
65-69	39,538	36,023	28,753	37,797	37,355	179,466
70-74	27,032	22,602	17,027	23,773	22,964	113,398
75-79	13,722	12,786	10,052	13,659	12,364	62,583
80+	5,712	5,407	3,844	5,536	4,881	25,380
All ages	2,504,165	2,371,363	1,660,491	2,486,734	2,248,262	11,271,005
Ages 5+	2,088,093	1,974,175	1,386,793	2,075,942	1,881,875	9,406,878

SOURCE: As for table 25.A.

data is required. But if the number of years in the intercensal interval is not an integral multiple of the number of years in the age group, adjustment procedures like those discussed in the preceding section will be required.

An illustration of the use of census survival ratios for the estimation of net migration by age is given in tables 25 to 28. Tables 25.A and 25.B give the age distribution of the male population in Korea in 1930 and 1935 by arbitrarily formed hypothetical groups of enumerated provincial data. In table 26, the five-year census survival ratios are shown for each age cohort. Table 27 shows the expected population in 1935 obtained by applying age-specific census survival ratios to the 1930 population (table 25.A) of the hypothetical regions. Finally, table 28 gives estimates of net migration obtained by subtracting the expected population from the population enumerated in 1935. These computational procedures for estimating net migration are essentially the same as those given in table 24 except that census survival ratios are used instead of life table survival ratios.

It should be noted that in table 28, the sum of net migration balances of all areal units is zero for each age

TABLE 26. MALE POPULATION OF KOREA, BY AGE, 1930 AND 1935, AND CENSUS SURVIVAL RATIO, 1930-1935

Age in 1930	Enumerated population, 1930	Age in 1935	Enumerated population, 1935	Census survival ratios, 1930-1935
0-4	1,661,240	5-9	1,478,064	0.88973538
5-9	1,361,625	10-14	1,301,810	0.95607087
10-14	1,153,608	15-19	1,080,314	0.93646542
15-19	1,058,199	20-24	959,748	0.90696362
20-24	860,573	25-29	811,545	0.94302866
25-29	692,154	30-34	652,568	0.94280753
30-34	706,726	35-39	669,307	0.94705303
35-39	618,681	40-44	572,611	0.92553513
40-44	548,241	45-49	508,714	0.92790215
45-49	460,921	50-54	409,120	0.88761415
50-54	379,603	55-59	334,045	0.87998514
55-59	305,594	60-64	248,205	0.81220508
60-64	235,138	65-69	179,466	0.76323691
65-69	171,553	70-74	113,398	0.66100855
70-74	114,519	75-79	62,583	0.54648574
75+	70,514	80+	25,380	0.35992852
All ages	10,398,889	Total 5+	9,406,878	

SOURCE: Computed from tables 25.A and 25.B.

Note: Census survival ratios are carried to eight decimal places only because of the correction procedure used for table 31.

TABLE 27. EXPECTED MALE POPULATION OF KOREA, BY AGE AND HYPOTHETICAL REGIONS, 1935

Age	Hypothetical "regions" A	B	C	D	E	Total
5-9	344,631	307,518	215,844	324,123	285,948	1,478,064
10-14	305,862	268,187	185,446	284,951	257,364	1,301,810
15-19	249,890	223,429	155,569	241,346	210,080	1,080,314
20-24	211,498	199,005	142,808	209,793	196,644	959,748
25-29	172,734	171,127	124,225	182,853	160,606	811,545
30-34	137,139	142,426	97,715	150,011	125,277	652,568
35-39	149,702	138,442	99,126	151,366	130,671	669,307
40-44	127,778	118,869	82,695	129,409	113,860	572,611
45-49	117,882	104,738	73,520	108,617	103,957	508,714
50-54	88,888	87,057	63,085	89,899	80,191	409,120
55-59	73,500	69,594	52,646	73,548	64,757	334,045
60-64	55,942	51,470	37,604	53,029	50,160	248,205
65-69	39,615	35,699	28,514	38,193	37,409	179,466
70-74	27,263	22,145	17,203	23,815	22,972	113,398
75-79	13,737	12,286	10,378	13,705	12,477	62,583
80+	5,461	5,205	4,178	5,617	4,919	25,380
TOTAL, ages 5+	2,121,558	1,957,197	1,390,556	2,080,275	1,857,292	9,406,878

SOURCE: Computed from tables 25.A and 26.

TABLE 28. NET INTERREGIONAL MIGRATION OF MALE POPULATION OF KOREA, BY AGE AND HYPOTHETICAL REGIONS, 1930-1935

Age	Hypothetical "regions" A	B	C	D	E
5-9	+1,564	+67	−6,008	−642	+5,019
10-14	−5,407	+1,553	+1,110	−691	+3,435
15-19	−11,554	+3,204	+1,859	+95	+6,396
20-24	−13,700	+8,205	+3,169	+6,188	−3,862
25-29	−4,880	+3,597	−346	+255	+1,374
30-34	−622	+611	−1,258	−2,059	+3,328
35-39	+95	−3	−1,195	−1,598	+2,701
40-44	+933	−662	−553	−1,591	+1,873
45-49	+695	−795	−590	−1,316	+2,006
50-54	−128	−452	+123	−1,203	+1,660
55-59	−99	−133	+257	−753	+728
60-64	−254	+303	+266	−453	+138
65-69	−113	+324	+239	−396	−54
70-74	−231	+457	−176	−42	−8
75-79	−15	+500	−326	−36	−113
80+	+251	+202	−334	−81	−38
TOTAL, ages 5+	−33,465	+16,978	−3,763	−4,333	+24,583

SOURCE: Computed from tables 25.B and 27.

group. This is always true whatever the nature of error in the age data or survival ratios and is one of the features that distinguishes the Census Survival Ratio method from the Life Table Survival Ratio method.

The CSR method is such that it tends to correct for systematic errors in the age data and thus to compensate for some of the effects of such errors. The age group 0-4 years, for example, may be disproportionately undernumerated. It often happens that this cohort is better enumerated in a later census (say, ten years later when the cohort is aged 10-14 years) and the number is found to be larger than would be expected on the basis of any reasonable estimate of change due to mortality. (CSRs for this cohort sometimes have values greater than unity.) Such ratios do not give accurate measures of survivorship, but they do tend to incorporate net census error in the expected population and to that extent give a better estimate of net migration than would a life table ratio which "expects" no change except that due to mortality. These differentials in the completeness of enumeration of a cohort at successive censuses cause CSRs to fluctuate somewhat rather than to follow the smoothly descending age-pattern characteristic of LTSRs. This feature of the CSR method in one of its advantages over the LTSR method.

But the CSR method has certain disadvantages. These become apparent when one examines the basic assumptions that are implicit in it:

(1) That the national population is closed, i.e., entered only by births and left only by deaths, and therefore is not affected by external migration;

(2) That the specific mortality rates are the same for each areal unit as for the nation;

(3) That the ratio of the degree of "completeness" of enumeration in any age-sex group in each areal unit (i.e., the proportion that any age-sex group bears to the true population) to that of the nation is the same for the same cohort in both censuses.

The accuracy of estimates of net migration will be affected by the extent to which these assumptions are met. Their validity should therefore be examined, as indicated in the following paragraphs.

Assumption (1)

An effort should be made to determine whether the native population was reasonably closed for the period under consideration. Some countries have fairly good data on international migration. These are useful for assessing the amount of international migration and its effect on national and regional population growth. Alternatively, data on country of birth will be helpful. In any case, it is essential to make an assessment of the importance of international migration and to adjust for it to approximate closure before census survival ratios are computed.

Assumption (2)

The matter of geographical uniformity in mortality should also be examined. In countries where the general mortality level is high, there is likely to be considerable variation in the mortality of component areas. The assumption of mortality equality may therefore be seriously violated; and if migration estimates are not corrected for regional differences in mortality, errors will be introduced. The error in the net migration estimates is equal to $\Delta S_i \cdot p_{i,x,t}$ if the estimates are based on the Forward CSR method, where ΔS_i denotes the difference between the survival ratio for the nation and that of the particular areal unit. The error in the net migration rate obtained by the same method, and using the population at the first census as a base, is therefore equal to ΔS_i. An idea of the extent of this error for an intercensal period of ten years is illustrated by the examples in table 29. Thus, if the national expectation of life at birth is 40 years, and if the area expectation of life at birth is 45 years, the error in the net migration rate stemming from this difference is 1.2 per cent for the age group 10-14 years; 1.8 per cent for the age group 20-24 years; 2.4 per cent for the age group 30-34 years, and so on.

As will be shown in the next chapter, there is an advantage in using the population at the second census as a base to calculating the rate of net migration by the Forward CSR method, since the error due to differences in degree of enumeration is cancelled; and the migration rate is thereby freed from enumeration errors.[4] However, the error in the migration rate based on the Forward CSR method, uncorrected for regional differences in mortality and using the population at the second census as a base, is equal to:

$$\Delta S_i \cdot \frac{p_{i,x,t}}{p_{i,x+n,t+n}}$$

As this formula indicates, the error in the migration rate for a given area varies systematically with the magnitude and direction of net migration.

Absence of information on areal differences in mortality is a major difficulty in making such adjustments. Where there is some idea about existing differences, adjustments can be done. The technique of adjustment is illustrated in tables 30-34, where the Korean provincial population age data from tables 25.A and 25.B are combined into five purely hypothetical "regions", each "region" being assigned arbitrarily an expectation of life. The differences between the assigned regional values and the national level of 45 years[5] are as follows:

Regions	e_0^0	*Deviation from* $e_0^0 = 45$
A	40.0	−5.0
B	42.5	−2.5
C	45.0	0.0
D	47.5	+2.5
E	50.0	+5.0

Corrections factors for each "region" were then calculated by dividing the survival ratios of the appropriate model life tables[6] by those that correspond to a life expectancy of 45 years. The correction factors obtained in table 30 were then applied to the census survival ratios of table 26 to obtain the set of adjusted ratios given in table 31. Application of these adjusted ratios to the 1930 population of the "regions" yields expected numbers for 1935 as shown in table 32. The sum of the expected populations for all sub-areas given in table 32 would be equal to the total national population only if the national survival ratios were a weighted average of net survival ratios of sub-areas. This condition would not ordinarily be met, and hence it is necessary to force redistributions by prorating differences, age group by age group. Finally, estimates of net migration were derived by subtracting the expected population of table 33 from the enumerated population in table 25.B. Table 34 gives the estimates of net migration for regions, the algebraic sum of which yields zero age by age.

Assumption (3)

It will rarely be possible to make a direct determination of the validity of the third assumption concerning areal variations in the relative undercount or overcount of age cohorts. The built-in mechanism that corrects for enumeration error will be effective only to the degree that this assumption holds. Moreover, even if this

[4] K. C. Zachariah, "A note on the Census Survival Ratio method of estimating net migration", *Journal of the American Statistical Association* (Washington, D.C.), March 1962, pp. 175-183.

[5] The national e_0^0 of 45 for male is approximately equal to the level given by Yun Kim after investigation of available life tables for the 1930s in his Ph.D. dissertation cited above in table 25.A.

[6] See *Manual III: Methods of Population Projections by Sex and Age.*

TABLE 29. DIFFERENCE IN SURVIVAL RATIOS, BY AGE, NATIONAL EXPECTATION OF LIFE AT BIRTH, AND THE DIFFERENCE IN EXPECTATION OF LIFE AT BIRTH BETWEEN COMPONENT AREAS AND THE NATION

	National $e_0^0 = 40$				National $e_0^0 = 50$				National $e_0^0 = 60.4$				National $e_0^0 = 65.8$			
	Area deviations				*Area deviations*				*Area deviations*				*Area deviations*			
Age	−10	−5	+5	+10	−10	−5	+5	+10.4	−10.4	−5.4	+5.4	+9.8	−10.8	−5.4	+4.4	+8.1
10-14	+.0295	+.0138	−.0119	−.0221	+.0221	+.0102	−.0095	−.0177	+.0177	+.0082	−.0072	−.0141	+.0153	+.0072	−.0069	−.0135
20-24	+.0426	+.0198	−.0175	−.0324	+.0324	+.0149	−.0140	−.0262	+.0262	+.0122	−.0108	−.0206	+.0229	+.0108	−.0099	−.0181
30-34	+.0632	+.0282	−.0237	−.0428	+.0428	+.0191	−.0166	−.0305	+.0305	+.0139	−.0116	−.0215	+.0255	+.0116	−.0099	−.0187
50-54	+.0981	+.0453	−.0387	−.0702	+.0702	+.0315	−.0280	−.0512	+.0512	+.0232	−.0191	−.0360	+.0423	+.0191	−.0169	−.0409

SOURCE: Adapted from K. C. Zachariah, "A note on the Census Survival Ratio method of estimating net migration", *Journal of the American Statistical Association*, vol. 57, March, 1962, p. 182.

TABLE 30. CORRECTION FACTORS FOR MORTALITY ADJUSTMENTS OF CENSUS SURVIVAL RATIOS FOR THE MALE POPULATION OF KOREA, 1930-1935

Age	Deviations from $e_0^0 = 45$			
	−5.0	−2.5	+2.5	+5.0
0-4	0.9796	0.9904	1.0088	1.0170
5-9	0.9939	0.9971	1.0027	1.0050
10-14	0.9945	0.9974	1.0024	1.0045
15-19	0.9929	0.9966	1.0032	1.0062
20-24	0.9912	0.9957	1.0040	1.0076
25-29	0.9898	0.9951	1.0044	1.0084
30-34	0.9881	0.9942	1.0050	1.0097
35-39	0.9855	0.9931	1,0059	1.0113
40-44	0.9827	0.9920	1.0071	1.0134
45-49	0.9797	0.9904	1.0083	1.0157
50-54	0.9760	0.9886	1.0099	1.0189
55-59	0.9718	0.9867	1.0117	1.0224
60-64	0.9667	0.9841	1.0143	1.0274
65-69	0.9584	0.9798	1.0183	1.0353
70-74	0.9468	0.9739	1.0240	1.0463
75+	0.9360	0.9688	1.0285	1.0549

SOURCE: Computed from *Manual III: Methods for Population Projections by Sex and Age* (United Nations publication, Sales No.: 56.XIII.3), table V, pp. 80-81. Correction factors are the ratios of the survival ratios (as denoted by P_x in the Source). For example, column (1) = (P_x corresponding to $e_0^0 = 40/P_x$ corresponding to $e_0^0 = 45$) etc.

TABLE 31. CENSUS SURVIVAL RATIOS FOR THE MALE POPULATION OF KOREA ADJUSTED FOR MORTALITY DIFFERENCES, BY AGE AND HYPOTHETICAL REGIONS, 1930-1935

Age	Hypothetical "regions"				
	A	B	C	D	E
0-4 to 5-9	.87158478	.88119392	.88973538	.89756505	.90486088
5-9 to 10-14	.95023884	.95329826	.95607087	.95865226	.96085122
10-14 to 15-19	.93131486	.93403061	.93646542	.93871294	.94067951
15-19 to 20-24	.90052418	.90387994	.90696362	.90986590	.91258679
20-24 to 25-29	.93473001	.93897364	.94302866	.94680077	.95019568
25-29 to 30-34	.93319089	.93818777	.94280753	.94695588	.95072711
30-34 to 35-39	.93578310	.94156012	.94705303	.95178830	.95623944
35-39 to 40-44	.91211487	.91914894	.92553513	.93099579	.93599368
40-44 to 45-49	.91184944	.92047893	.92790215	.93449026	.94033604
45-49 to 50-54	.86959558	.87909305	.88761415	.89498135	.90154969
50-54 to 55-59	.85886550	.86995331	.87998514	.88869699	.89661686
55-59 to 60-64	.78930090	.80140275	.81220508	.82170788	.83039847
60-64 to 65-69	.73782112	.75110144	.76323691	.77415120	.78414960
65-69 to 70-74	.63351059	.64765618	.66100855	.67310501	.68434215
70-74 to 75-79	.51741270	.53222246	.54648574	.55960140	.57178803
75+ to 80+	.33689309	.34869875	.35992852	.37018648	.37968860

SOURCE: Computed from tables 26 and 30. See Note to table 26.

assumption does hold, the estimated amount of net migration in each age group will differ from the true amount by the same proportion that the enumerated population at the second census differs from the true population of the given age-sex group. If the age distribution is defective only because of mis-statement of age and not because of under-enumeration or over-enumeration, some of the error is eliminated by combining the estimates into broad age groups. If the age distribution is defective because of under-enumeration or over-enumeration of the general population, the total amount of net migration (all ages combined) will tend to be underestimated or overestimated, according to the direction of the enumeration error.

It should be noted here that if the estimate of net migration has the same relative error as the population at the second census (which can happen even if the third assumption is valid), a *rate* of migration based upon the population at the second census will be free of enumeration error.[7]

It should also be noted that the forward CSR method yields an estimate of net migration only for survivors of the initial age cohort. If the number of deaths to in-migrants differs from the number of deaths to out-migrants, the estimated net migration will differ from

[7] See K. C. Zachariah, op.cit., pp. 175-183.

TABLE 32. FIRST APPROXIMATION TO ESTIMATES OF THE EXPECTED MALE POPULATION OF KOREA, BY AGE AND HYPOTHETICAL REGIONS, 1935

Age	Hypothetical "regions"					Total
	A	B	C	D	E	
5-9	337,600	304,567	215,844	326,976	290,809	1,475,796
10-14	303,996	267,410	185,446	285,721	258,651	1,301,224
15-19	248,515	222,848	155,569	241,926	211,025	1,079,883
20-24	209,997	198,328	142,808	210,463	197,863	959,459
25-29	171,213	170,391	124,226	183,585	161,827	811,242
30-34	135,740	141,729	97,714	150,671	126,329	652,183
35-39	147,921	137,639	99,126	152,122	131,938	668,746
40-44	125,925	118,049	82,694	130,173	115,147	571,988
45-49	115,844	103,899	73,519	109,388	105,349	507,999
50-54	87,084	86,221	63,086	90,645	81,450	408,486
55-59	71,736	68,801	52,646	74,276	65,981	333,440
60-64	54,364	50,786	37,604	53,650	51,284	247,688
65-69	38,330	35,131	28,514	38,739	38,435	179,149
70-74	26,129	21,697	17,203	24,251	23,783	113,063
75-79	13,006	11,965	10,378	14,033	13,055	62,437
80+..........	5,112	5,043	4,178	5,777	5,189	25,299
TOTAL, ages 5+.	2,092,512	1,944,504	1,390,555	2,092,396	1,878,115	9,398,082

SOURCE: Tables 25.A and 31. Adjustment is made for mortality differences only.

TABLE 33. FINAL ESTIMATES OF THE EXPECTED MALE POPULATION OF KOREA, BY AGE AND HYPOTHETICAL REGIONS, 1935

Age	Hypothetical "regions"					Total
	A	B	C	D	E	
5-9	338,208	305,115	215,844	327,565	291,332	1,478,064
10-14	304,156	267,550	185,446	285,871	258,787	1,301,810
15-19	248,631	222,952	155,569	242,039	211,123	1,080,314
20-24	210,071	198,398	142,808	210,538	197,933	959,748
25-29	171,289	170,466	124,226	183,666	161,898	811,545
30-34	135,834	141,827	97,714	150,776	126,417	652,568
35-39	148,067	137,774	99,126	152,272	132,068	669,307
40-44	126,085	118,199	82,694	130,339	115,294	572,611
45-49	116,035	104,070	73,519	109,568	105,522	508,714
50-54	87,244	86,379	63,086	90,811	81,600	409,120
55-59	71,891	68,949	52,646	74,436	66,123	334,045
60-64	54,498	50,911	37,604	53,782	51,410	248,205
65-69	38,411	35,205	28,514	38,820	38,516	179,466
70-74	26,220	21,773	17,203	24,336	23,866	113,398
75-79	13,042	11,999	10,378	14,072	13,092	62,583
80+	5,132	5,062	4,178	5,799	5,209	25,380
TOTAL, ages 5+.	2,094,814	1,946,629	1,390,555	2,094,690	1,880,190	9,406,878

SOURCE: Tables 32 and 25.B.

Note: Differences between total columns in the two tables are prorated age group by age group for regions A, B, D and E to achieve equality between the observed total population and the expected total population. The data for region C are not prorated because it is assumed that the survival ratios for this region are the same as those of the nation.

the true amount. This follows from the fact that $(1-S)p_{i,t}$ yields the deaths among all those who were enumerated in the first census, including those who migrated out of the area. Moreover, the enumerated population $p_{i,t+n}$ does not include those who migrated to the area after the first census and who died there before the second census. If we denote by D_O the number of deaths among persons who moved out of the area during the intercensal interval and by D_I the number of deaths among those who migrated into the area during the same period, an expression that includes the migration of such persons is given by:

$$\text{Net } M_i = (p_{i,t+n} - Sp_{i,t}) + (D_I - D_O) \qquad (14)$$

If $D_I = D_O$, the error in the estimates from this source is zero. If the death rate is very low, as among those who are in the age group 10-14 or 15-19, D_I and D_O will both be small, and hence $(D_I - D_O)$ will be negligible. Correspondingly, if all migrations (both in and out) took place near the end of the intercensal period, the error will be

TABLE 34. NET INTERREGIONAL MIGRATION OF MALE POPULATION OF KOREA, BY AGE AND HYPOTHETICAL REGIONS, 1930-1935

Age	Hypothetical "regions" A	B	C	D	E
5-9	7,987	2,470	−6,008	−4,084	−365
10-14	−3,701	2,190	1,110	−1,611	2,012
15-19	−10,295	3,681	1,859	−598	5,353
20-24	−12,273	8,812	3,169	5,443	−5,151
25-29	−3,435	4,258	−347	−558	82
30-34	683	1,210	−1,257	−2,824	2,188
35-39	1,730	665	−1,195	−2,504	1,304
40-44	2,626	8	−552	−2,521	439
45-49	2,542	−127	−589	−2,267	441
50-54	1,516	226	122	−2,115	251
55-59	1,510	512	257	−1,641	−638
60-64	1,190	862	266	−1,206	−1,112
65-69	1,127	818	239	−1,023	−1,161
70-74	812	829	−176	−563	−902
75-79	680	787	−326	−413	−728
80+	580	345	−334	−263	−328
TOTAL, all ages ...	−6,721	25,546	−3,762	−18,748	1,685

SOURCE: Tables 25.B and 33.

small. The error will be greatest when the mortality rates are high, when the amounts of in-migration and out-migration differ considerably and when most of the migration took place near the beginning of the intercensal period. Use of the average CSR method may be indicated when it is specifically desired to include an allowance for the death factor in the measure of net migration.

In summary, it appears that Census Survival Ratio methods furnish a simple and convenient means of estimating net migration where area-specific age data are tabulated in the census. It should, however, be borne in mind that the estimates will be valid only if the assumptions discussed above hold reasonably well for the country and for the period. The validity of all the assumptions (particularly about the absence of international migration and equality of areal mortality levels) should be carefully examined. If the facts are not consistent with these assumptions, every possible attempt should be made to adjust the migration estimates.

It may sometimes happen that indirect evidence shows that the mortality differences are large; but in the absence of any quantitative information about these differences, it is not feasible to make any adjustments. If, however, an independent estimate of the amount of net migration at all ages is obtained by the Vital Statistics method or the Place-of-Birth method, the results of the CSR method are useful for distributing the total amount of net migration by age.

Net migration of children

The Census Survival Ratio method cannot give estimates of net migration for persons born during the intercensal interval. This gap may be filled by various methods. If the birth registration is considered to be complete and numbers of births are available by areal units, these can be used to calculate survival ratios and for computing estimates of net migration. Thus, if data by quinquennial age groups are available from a census taken on 1 January 1960, after an intercensal interval of ten years, survival ratios for quinquennial age groups are given by:

$$S_1 = \frac{\text{National population 0-4 years old in 1960}}{\text{National births during 1955-1959}} \tag{15}$$

$$S_2 = \frac{\text{National population 5-9 years old in 1960}}{\text{National births during 1950-1954}} \tag{16}$$

An estimate of net migration for persons 0-4 years old in 1960 in the i^{th} area is given by:

$$\text{Net } {}_5M_{0,i} = {}_5p_{0,i} - S_1 \cdot B_i\,(1955\text{-}1959) \tag{17}$$

and that for persons aged 5-9 years is given by:

$$\text{Net } {}_5M_{5,i} = {}_5p_{5,i} - S_2 \cdot B_i\,(1950\text{-}1954) \tag{18}$$

These estimates, like those for the older cohorts, have the property that their total for all areas will automatically be zero.

If reliable birth statistics are not available, the following approximate method, which uses area-specific child-woman ratios, derived from the second census, may be applied.[8] If the ratios of children aged 0-4 to women aged 15-44 and of children aged 5-9 to women aged 20-49 are denoted by CWR_0 and CWR_5 respectively, then estimates of net migration for the age groups 0-4 (denoted by Net ${}_5M_{0,i}$) and 5-9 (denoted by Net ${}_5M_{5,i}$) are given by:

$$\text{Net } {}_5M_{0,i} = \tfrac{1}{4}\,\text{CWR}_0\ \text{Net } {}_{30}M^{(f)}_{15,i} \tag{19}$$

$$\text{Net } {}_5M_{5,i} = \tfrac{1}{4}\,\text{CWR}_5\ \text{Net } {}_{30}M^{(f)}_{20,i} \tag{20}$$

[8] This method was devised by Everett S. Lee in *Population Redistribution and Economic Growth, United States, 1870-1950; I. Methodological Considerations and Reference Tables* (Philadelphia, The American Philosophical Society, 1957), p. 65.

where Net ${}_{30}M^{(f)}_{15,i}$ and Net ${}_{30}M^{(f)}_{20,i}$ are the area estimates of net migration for females aged 15-44 and 20-49 respectively. If we assume that the flow of migration was even and fertility ratios constant, then one fourth of the younger and three fourths of the older children would have been born before their mothers migrated. The sum of these net migration estimates for all areas will not necessarily be zero. A zero balance can be obtained by raising the amounts of net loss and lowering the amounts of net gain, or *vice versa*, in proportion to the unadjusted estimates.

Special cases

The detail in which statistics are tabulated may vary from one census to the next. Thus, data by age and sex may not be available from both censuses for all the areas for which it is desired to derive estimates of net internal migration. If the population has been classified by age for the country as a whole at the two censuses and age detail for component areas is available only at the second census, it is possible to derive estimates of net migration, by areas, for ages n years and above combined. This can be done by applying age-specific reverse survival ratios to the age distributions of the second census, summing the expected numbers, and subtracting the observed population at the first census from the expected total. The estimates of total migration so obtained can be made comparable with those that would have resulted from the Forward Survival Ratio method if we multiply them by the over-all forward survival ratio:

$$ {}_nS_{0+} = \frac{P_{n+,t+n}}{P_{0+,t}} \qquad (21) $$

If the age classification is available only at the second census both for the country as a whole and for component areas, it is still possible to derive rough estimates of net migration by the use of an alternative short method. An over-all forward survival ratio for the country as a whole can be applied to the population in the first census for the areas for which the migration estimate is required, thus obtaining an estimate of the "expected" population of that area n years old and over at the second census and deriving an estimate of net migration by subtraction.[9] Symbolically:

$$ \text{Net } M_i(n+) = p_{i,n+,t+n} - \frac{P_{n+,t+n}}{P_{0+,t}} \cdot p_{i,0+,t} \qquad (22) $$

Comparison of indirect measures with other measures

The following types of sources for obtaining estimates of net migration for fixed intervals have been described:

(1) Place of residence at a fixed past date

(2) Duration of residence, by place of last residence

(3) Place of birth statistics (POB)

(4) Vital statistics (VS)

(5) Survival ratios (LTSR or CSR)

The first two yield "direct" measures of net migration. As indicated in previous discussion, the definition of migrants differs slightly as between the two. In (1) a migrant is a person whose place of residence at the time of enumeration differs from his place of residence at the beginning of the interval. In (2) a migrant is a person who has changed his place of residence at least once during the interval.

Sources (3), (4) and (5) yield "indirect" measures of net migration. In all three, the implied definition of a migrant is the same as in (1). Measures derived from (1) are therefore more compatible with the indirect measures than are those derived from (2).

Since the last three methods are designed to measure the same phenomenon, they should provide the same estimate for a given interval. In practice, however, the results obtained differ somewhat and there is need to look into the possible causes of discrepancy among them and to make an assessment of the relative efficiency of each.

The VS method is unique in the sense that it is the only method which measures the balance of all migrations made during the interval. A person who came into the area during the intercensal period and died before the end of the period would be counted in the measure of net migration. Similarly, a person who moved out during the intercensal period and died outside before the second census would be counted in the estimate of net migration. But, as shown earlier, there are sources of errors in this estimate which may be serious in a situation where the errors in the basic data are large. The major points to note in this connexion are:

(1) For the same relative error, the amounts of error in the population data are more important than those in the vital statistics;

(2) Accurate vital statistics in combination with defective population data are likely to give imprecise migration estimates.

It should also be noted that, under prevailing conditions, this method yields estimates only for all ages combined. Detail by age, an important consideration in migration analysis, is not ordinarily obtainable.

The Forward CSR method (the one that is generally preferred) measures net migration, by age, for persons alive at the end of the intercensal period. This measure does not take account of the migration of persons who died during that period. Only in the unlikely event that deaths among out-migrants equal deaths among in-migrants during the intercensal period does the Forward CSR method give net migration among persons alive at the beginning of the intercensal period.[10] Since the number of deaths is likely to be larger in the larger of the two components of net migration (in-migration and out-migration), CSR estimates obtained by the forward method will generally be smaller than those obtained by the VS method.

[9] For examples of the application, of short methods, see D. J. Bogue and K. C. Zachariah, "Urbanization and migration in India" in Roy Turner, ed., *India's Urban Future* (Berkeley, University of California Press, 1962), pp. 27-54.

[10] Although the average CSR method is designed to take account of migrant deaths, it involves assumptions about the distribution of migration and deaths that may not be met. With the VS method, this problem does not arise.

It is difficult to make a general statement, valid in all situations, regarding the relative accuracy of all-ages estimates from the two methods. Estimates from the CSR method have certain theoretical advantages. For a country with a closed population, with a low over-all mortality level, with negligible regional variation in mortality conditions and reasonably good age data, this method is probably preferable to the VS method, if for no other reason than that it yields information by age.[11] On the other hand, if the regional variation in mortality conditions is great and if the age data are highly deficient, the theoretical advantages of the method will not stand up against the errors due to mis-statement of age, under-enumeration or over-enumeration, and most important of all, the error arising from mortality differences among geographic areas. But under these conditions, the vital statistics are also generally poor and the CSR method may be the only one open to the research worker. If no independent estimates are available, one should at least examine the internal consistency of the estimates, particularly the smoothness of the age curve of migration rates, and attempt to evaluate the influence of regional differences in mortality levels.

As between the CSR method and the LTSR method, the former is generally preferable. The reasons have already been stated. However, the situation can arise in which the LTSR method is preferable. For example, if migration estimates are required for only one or two small areas in a country (e.g., one city) and the mortality level is known to be different from that of the country as a whole or if the national population is not sufficiently closed and no satisfactory adjustment can be made for international migration, the LTSR method may be indicated. If a life table is available for the area, it should be used; otherwise one can utilize model life tables. If the age data are defective, it is desirable to smooth them before the life table survival ratios are applied. The range of the smoothing formula should be relatively small, covering not more than 10 years of age at a time (if single-year-age data are available) or not more than three five-year age groups (if only five-year age groups are available).[12]

Not much is known about the validity of responses to the question on birth-place. It seems likely that in a country where the census data and vital statistics are reasonably good and where the assumptions about international migration and mortality differences are satisfactorily met, net migration estimates by the VS method or the CSR method are more accurate than estimates by the POB method (procedure 1). But in some situations, the latter method is definitely to be preferred. Since the VS and the CSR methods have already been compared, we limit the present comparison to the CSR and POB methods. Estimates by these two methods may be written as follows:

$$\text{CSR Method: Net } M'_{CSR} = p_{t+n} - S \cdot p_t \tag{23}$$

$$\text{POB Method: Net } M'_{POB} = (I_{t+n} - O_{t+n}) - S(I_t - O_t) \tag{24}$$

It can easily be shown that if the census figures are accurate and if the survival ratio is correctly estimated, Net M'_{CSR} and Net M'_{POB} will be identical. If, however, the above conditions are not met, there will be appreciable differences between the two estimates. For any areal unit, p_t and p_{t+n} will be much larger than $(I_{t+n} - O_{t+n})$ and $(I_t - O_t)$, and Net M'_{CSR} will be the difference between much larger numbers than will Net M'_{POB}. It follows that, for a given percentage error in the survival ratio, the amount of error in Net M'_{CSR} is likely to be much larger than that in Net M'_{POB}. Where regional mortality differences are large and age data are seriously defective, there is additional reason to expect the POB method to give more valid estimates of net migration. This is likely to be true even if the birth-place data are not tabulated by age at all and if only approximate survival ratios are applied to lifetime in-migrants and lifetime out-migrants. If the birth-place data are available also by age for each lifetime stream, the possibility of eliminating the error due to mortality differences is an important consideration operating in favour of the POB method (procedure 3).

There are, however, other considerations which favour the CSR method. In the first place, the survivorship of the populations involved in equation (23) differs from that in equation (24) and unless the birth-place data are available by age, the S in the former is likely to be more accurate than the S in the latter. Secondly, the percentage errors (particularly errors due to under-reporting) in I and O are likely to be greater than those in p_t and p_{t+n} and there is, therefore, a tendency for birth-place data to underestimate migration. These considerations do not invalidate the main points made above, but more information about the validity of the responses on birth-place is necessary before a definitive judgement about the relative accuracy of the POB method can be made for countries where the mortality level is relatively low and census and registration data are reasonably good. Where feasible, evaluation of the two approaches is aided by preparing estimates according to both methods and comparing the results.[13]

[11] For a comparison of results obtained by the VS and Forward CSR methods, see Hope T. Eldridge, *Net Intercensal Migration for States and Geographic Divisions of the United States, 1950-1960: Methodological and Substantive Aspects*, Analytical and Technical Report No. 5 (Population Studies Center, University of Pennsylvania, Philadelphia, 1965), chapter VII. For a comparison of the VS and Average CSR methods, see C. Horace Hamilton, "Effect of census errors on the measurement of net migration", *Demography* (Chicago), vol. 3, No. 2, 1966, pp. 393-415.

[12] A simple moving average formula with equal weights for ten-year coverage may be used for single-year-age data; and a 3-point formula such as [¼, ½, ¼] may be used for five-year age groups. For a detailed discussion of the reasons for and the assumptions involved in such smoothing procedures, see K. C. Zachariah, op. cit., pp. 149-151.

[13] Such comparisons for the United States of America, 1950-1960, are made in Eldridge and Kim, op. cit., chapter III.