Title

A COMPUTERIZED MODEL TO ESTIMATE CURRENT NET MIGRATION RATES BY RACE, SEX, AND AGE

Author

Schroeder, E.C.

Publication Date

1981-06-01

LBL-12917
UC-13

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

Physics, Computer Science & Mathematics Division

A COMPUTERIZED MODEL TO ESTIMATE CURRENT NET MIGRATION RATES BY RACE, SEX, AND AGE

Esther C. Schroeder

June 1981

Prepared for the U.S. Department of Energy under Contract W-7405-ENG-48

DISCLAIMER

A Computerized Model

to Estimate

Current Net Migration Rates

by Race, Sex, and Age

June 1981

Esther C. Schroeder
Computer Science and Applied Mathematics
Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

This work was supported by the U. S. Department of Energy
under Contract number W-7405-ENG-48.

Table of Contents

I. Introduction

The population, labor force and unemployment projections model developed at Lawrence Berkeley Laboratory as part of the Labor Market Projections Model (LMPM) uses a cohort-component model to project population at the local level (Schroeder, 1980). Currently, this model is projecting from 1970 to 1982. This is done in two steps - first by projecting to 1975 and then to 1980. The 1982 control totals are applied to the 1980 distribution to yield the 1982 population projections by race, sex and age. The model is designed to operate at various geographic levels - states, SMSAs, counties, and cities - with particular emphasis on the smaller areas. Understandably, the model is sensitive to the age-specific net migration rates used; particularly at the substate level. The most recent net migration rates available by race, sex, and age at the county level are the 1960-1970 rates as developed by Bowles, Beale, and Lee, (1975). However, the migration patterns in many geographic areas have changed considerably from the 1960's to the 1970's. Using the 1960-1970 net migration rates to project into the 1980's can be inappropriate.

The first attempt to update the 1960-1970 migration rates used the Plus-Minus Technique (Shryock, Siegal and Associates, 1973). By using estimated net migration totals for 1970-1975 as obtained from the Revenue Sharing data (Census, 1976), this method updates the 1960-1970 net migration rates by race, sex and age so that when applied to the 1970 population of the area, they yield the desired net migration total for 1970-1975. This method works satisfactorily only for those areas in which net migration in the 70's follows the same pattern as that in the 60's. However, in most areas, migration patterns have changed considerably since the 60's. In these cases, the Plus-Minus Technique does not work well or breaks down completely.

The Plus-Minus Technique stretches or compresses net migration rates for individual age cohorts so that a desired net migration total can be obtained when these rates are applied to the population in the base year. A plus factor is applied to all positive rates that stretches them if more in migration is desired, or compresses them if less in migration is desired. A minus factor acts in a similar manner on the the negative rates. The technique is inflexible in that individual migration rates are not allowed to change sign; instead, one scale factor is applied to all positive rates and another scale factor is applied to all negative rates. If the change in migration is so great that it is impossible to obtain the desired net migration total merely by compressing and stretching the individual migration rates, the technique causes one of the adjustment factors to have a minus sign. The larger positive rates could then become the larger negative rates or the larger negative rates could become the larger positive rates. In either case the pattern of the original distribution has been changed considerably.

To obtain better estimates of net migration in the 1970's, a flexible migration submodel was developed based on a methodology described in detail in earlier publications (Pittenger, 1974 and 1978). To estimate the current net migration rates of an area, the "character" of that area is used. Areas of a similar type, e.g., rural, central city, suburban, etc., usually follow similar migration patterns. A set of about a dozen migration patterns was developed that mimicked most net migration flows. Once a pattern was chosen for a particular area, it was scaled by using total net migration estimates for the area.

A thorough testing of the results of this submodel and the population projections in general can not be done until data from the 1980 Census is available. However, some measure of its performance was obtained by comparing the population projections to 1975 with the Bureau of the Census County Population Estimates for 1975 (Census, 1980a). Without exception, the population projections using this submodel were much closer to the Census Estimates than were the population projections using the Plus-Minus technique on the 1960-1970 net migration rates.

II. Implementation

A. Background

The accuracy of short term population projections at the local level is dependant upon the migration rates used. Since migration flows can have considerable impact not only on the population of an area but also on the age distribution of that population, errors in estimating migration are potentially much more serious than those in estimating fertility or mortality. However, the migration component is at the same time the most difficult to estimate.

Migration rates by race, sex and age at the local level usually are calculated by using vital statistics (births and deaths) data and two Censuses. By the time they are calculated, they are measuring flows that occured well in the past. Furthermore, these historical data are not always valid for describing subsequent migration. In particular, age-specific migration rates are suspect if overall net migration levels have shifted markedly in a positive or negative direction since the most recent census.

The purpose of this computerized submodel is to provide a flexible means of estimating current migration patterns. To obtain estimates of current migration flows and rates, a migration submodel was developed

1) to estimate the overall volume of current net migration in five year time intervals after 1970, and

2) to estimate the actual number of migrants by race, age and sex from 1970-1975 and from 1975-1980.

With the post censal data presently available, it is not difficult to solve the first problem. In order to solve the second problem, it is necessary to obtain the total net migration flows by race and sex. Since migration flows vary considerably, particularly by race, it is not sufficient to obtain just an overall volume of net migration.

Due to the number of different areas which are handled by this model, it is impossible to study each area individually. However, by analyzing certain key data items, it is possible to estimate what is probably happening in each area. To estimate the age breakdown of the net migration total for a race-sex group, the following procedure is used. The submodel determines the character of the particular area from past migration rates. Once the character has been determined, migration "patterns" are assigned to the area. These "patterns" are then calibrated to yield a desired net migration total. Since retirement migration can vary considerably within migration "patterns", it is handled separately. Thus the migration patterns used by the model are only defined up to age 65.

The following discussion will concentrate on how this model was implemented in the computerized population projections model in LMPM - first the migration patterns as developed for the population up to age 65, and then for the retirement aged population. It should be kept in mind that this model is used to prepare estimates for more than one thousand areas across the United States. The model relies on data sources that are available nationally. Although the user can input data if he desires, the model does not depend upon user supplied information.

The key to understanding age-specific net migration rates and modeling them flexibly is to disaggregate them into their directional components - in migration and out migration. Net migration patterns can be complicated as their structure reflects the combination of in and out migration. It is easier to understand and to model the in and out migration flows of a particular area. By varying the relative volume of the directional rates and the age at which the maxiumum rate occurs, it is possible to mimic changes in empirical net migration patterns.

Ideally, historical directional flow patterns should be used to calibrate the model. This is not possible however, as five year race/age/sex directional net migration data are available only for states and state economic areas. But the fact that the model will be used for a variety of geographical data means that the only generally available data that can be used for calibration are the intercensal (10-year) migration estimtes such as the Bowles data mentionned above. In summary, this means that an indirect calibration procedure is required.

This procedure is discussed in Sections B and C. Section B will give a brief rationale for the various steps taken in setting up the model. For more details regarding the underlying methodology, please refer to Pittenger's articles as cited in references (2) and (3). Section C describes how the model handles the population under 65. Section D covers the migration of the population 65 and over. Section E covers adjustments that are necessary for the special populations - the military and the college students. Section F explains how to combine the results of the three preceding sections to yield the estimated net migration rates for each race/sex group.

B. Defining Migration Patterns

Model migration patterns can be defined by specifying two factors - the height of the peak age-specific rate and the age at which this peak occurs. The goal was to be able to mimic most migration curves with a limited number of patterns. Patterns were defined for three amplitudes at the peak (short, tall and extreme), and for four timings of the peak (college, early, intermediate and late). The source for analyzing age-specific directional migration flows was data published for the State Economic Areas (SEAs) (Census, 1963 and 1972). SEA data are very useful for several reasons - they show migration rates over a five year time period, they are defined for the same geographic areas for two time periods, 1955-1960 and 1965-1970, and they represent a variety of demographic conditions; central city counties, suburban counties, growing and declining areas, etc. After studying graphs of these data for many areas, it becomes apparent that several patterns are recurrent.

Further study of the age-specific directional migration curves indicates that the downward slope of the curve after the peak seems to vary with the level of migration. Figure 1 shows the slope of the directional migration rate at ages 40-44 for male populations in selected SEAs plotted against the percent directional migration for the same cohort. The upper plot shows outmigration and the lower, inmigration. Both show that at ages 40-44, the slope of the directional migration rate becomes less steep with increasing migration. To preserve this relationship, migration rate patterns were developed for three different slope values, 0.09, 0.12 and 0.15. Since 12 patterns were defined for each slope, there were 36 patterns in all. They are shown in Table 1 (Pittenger, 1980). Within each of the three slope categories, the pattern definition does not change for those cohorts aged less than 15 years, or for those aged 35 and older. Thus the number that is printed under CS for these age cohorts is identical across all patterns within the slope category.

For simplication, within each slope, it was decided to define just one migration pattern for the age cohorts aged less than 15 years and for those aged 35 and over. Since most of the migration occurs in the age cohorts 15-34, the migration submodel is most concerned with estimating the migration flows in these cohorts. At the present, it does not seem necessary to distinguish different migration patterns within each slope value for the remaining age cohorts. Further study may indicate that there should be different definitions for those aged less than 15 or for those aged 35 and over.

II. Implementation

TABLE 1 – Model Directional Migration Rates by Age, Slope and Pattern Type

	Pattern Type											
Age Cohort	CS	CT	CX	ES	ET	EX	IS	IT	IX	LS	LT	LX
						Slope 0.09						
(0- 4)	.1490											
(5- 9)	.2450											
(10-14)	.1980											
(15-19)	.3800	.493	.620	.235	.235	.235	.220	.220	.220	.180	.180	.180
(20-24)	.4000	.400	.450	.400	.500	.620	.330	.475	.580	.280	.280	.280
(25-29)	.3260	.326	.326	.326	.385	.385	.330	.420	.520	.326	.450	.550
(30-34)	.2650	.265	.265	.265	.265	.265	.265	.290	.290	.265	.330	.330
(35-39)	.2150											
(40-44)	.1750											
(45-49)	.1420											
(50-54)	.1160											
(55-59)	.0940											
(60-64)	.0764											
						Slope 0.12						
(0- 4)	.1250											
(5- 9)	.1900											
(10-14)	.1440											
(15-19)	.3500	.498	.620	.175	.175	.175	.160	.160	.160	.125	.125	.125
(20-24)	.3770	.377	.425	.377	.450	.550	.290	.420	.520	.240	.240	.240
(25-29)	.2860	.286	.286	.286	.330	.330	.290	.370	.470	.286	.390	.480
(30-34)	.2170	.217	.217	.217	.217	.217	.217	.235	.235	.217	.255	.255
(35-39)	.1650											
(40-44)	.1250											
(45-49)	.0948											
(50-54)	.0719											
(55-59)	.0546											
(60-64)	.0414											
						Slope 0.15						
(0- 4)	.0715											
(5- 9)	.1040											
(10-14)	.0722											
(15-19)	.2200	.337	.425	.090	.090	.090	.080	.080	.080	.060	.060	.060
(20-24)	.2390	.239	.280	.239	.300	.540	.180	.275	.350	.120	.120	.120
(25-29)	.1690	.169	.169	.169	.200	.200	.180	.240	.310	.169	.250	.325
(30-34)	.1200	.120	.120	.120	.120	.120	.120	.130	.130	.120	.150	.150
(35-39)	.0848											
(40-44)	.0600											
(45-49)	.0425											
(50-54)	.0301											
(55-59)	.0213											
(60-64)	.0151											

C = College, E = Early, I = Intermediate, L = Late
S = Short, T = Tall, X = Extreme

C. Pattern Assignment

The preceeding section defined the migration patterns that are used by the migration submodel. This section explains how a particular in and out migration pattern is chosen for each race/sex group in a given area. The decision is based on post censal population data and the migration rates of that race/sex group in the area for the time period 1960-1970. Appendix A lists the various steps and equations used by the model. The following discussion may be easier to understand if the reader also refers to this Appendix.

C.1 Estimating Overall Net Migration

Before deciding which migration patterns should be used in a particular area, the overall volume of net migration must be determined. Since migration flows in any area are apt to be considerably different for each race-sex group, the following procedure is used for each race-sex group. The volume of migration for a race-sex group is calculated by comparing an estimate of the 1975 population by race and sex (Census, 1980a) with a 1975 population comprised of survivors of the 1970 population plus survivors of those born from 1970-1975 (equation 1 in Appendix A). The difference is the total net migration by race and sex. The ratio of this total net migration to the 1970 population survived to 1975 is the total net migration rate (equation 2 in Appendix A). These total net migration rates from 1970-1975 are thus defined for the four race-sex groups - white males, white females, nonwhite males and nonwhite females.

C.2 Estimating In- and Out-Migration Rates

In defining the directional migration flow patterns, considerable attention was paid to the rates at age cohort (40-44) which has useful properties as a point of origin for certain analytical procedures described below (Pittenger, 1978). To decide which pattern is applicable it is necessary to determine directional migration rates and slopes at this age cohort. Figure 2 shows a plot of net migration rates for the age cohort (40-44) vs. the total net migration rate for the race-sex group. It appears from this plot that the net migration rate at the age cohort (40-44) can be approximated by that of the total race-sex group, ie., the percent net migration at age (40-44) is very close to the percent net migration of the entire group in many areas.

II. Implementation

It is now necessary to go from the percent net migration at age (40-44) to the percent in- and out-migration at age (40-44). Figure 3 shows plots of the in-migration rate at age (40-44) vs. net migration at age (40-44). The upper graph is for males and the lower is for females. Although there is some scatter in the plots, it does appear that it is possible to use the net migration rate for the race-sex group to estimate the in-migration rate for age cohort (40-44). Table 2 (Pittenger, 1980) is a tabular form of the data plotted in Figure 3 and is what was used by the model. Thus by a table look-up using the net migration rate, it is possible to obtain the in-migration rate. The out-migration rate is then obtained using the identity Out = In-Net.

TABLE 2 - Assignment of In-Migration Rates for Ages (40-44) Given Net Migration for Ages (40-44)

Net Migration Rate	In-Migration Rate Male	In-Migration Rate Female
< 4%	10%	7%
-4% to -3%	10%	7.5%
-3% to -2%	10.5%	8%
-2% to -1%	11%	8.5%
-1% to 0%	12%	9.5%
0% to 1%	13%	10.5%
1% to 2%	14%	11.5%
2% to 3%	15%	12.5%
3% to 4%	17.5%	13.5%
4% to 5%	19%	14.5%
5% to 6%	Net + 15%	15.5%
6% to 7%	Net + 15%	16.5%
7% to 8%	Net + 15%	17.5%
8% to 9%	Net + 15%	18.5%
9% to 10%	Net + 15%	20%
10% to 11%	Net + 15%	22%
11% to 12%	Net + 15%	24%
12% to 13%	Net + 15%	26%
13% to 14%	Net + 15%	27.5%
14% to 15%	Net + 15%	29%
> 15%	Net + 15%	Net + 15%

C.3 Estimating the Slope

As was mentioned in Section B above, the slope of the directional migration rate at ages 40-44 tends to become less steep with increasing migration. This information was used in defining the directional migration rates. Table 3 (Pittenger, 1980) is a tabular form of the data plotted in Figure 1. With the estimated in- and out-migration rates at ages (40-44) as just obtained and the figures in Table 3, it is possible to assign a slope for the in- and out-migration pattern for each race-sex group.

TABLE 3 - Assignment of Slope Categories
Given Direction and Value of Migration Rates at Ages (40-44)

	MALE		FEMALE	
Slope	In-Mig	Out-Mig	In-Mig	Out-Mig
0.09	>.18	>.12	>.19	>.13
0.12	Otherwise	Otherwise	Otherwise	Otherwise
0.15	<.09	<.09	<.10	<.10

C.4 Chosing the Directional Migration Patterns

Once the slope has been determined for a race-sex group in a particular area, a decision must still be made on which of the twelve patterns within the slope grouping best describe the character of this area. The applicable pattern is identified by examining the 1960-1970 net migration rates for the age cohorts (15-19) through (35-39). Since these are the age cohorts where the majority of the migration occurs and where changes in inflection of net rate patterns are usually found, the differences between migration patterns are most evident in these age cohorts. Let the net migration ratio be defined as one plus the net migration rate (step 6 in Appendix A). Calculate the net migration ratio for age cohorts (15-19) through (35-39), ie., for age cohorts with indices 4 through 8. Table 4 shows data for white males in West Virginia for the decade 1960-1970 (Bowles et al., Part 3, page 64).

II. Implementation

Of the five age cohorts of interest, the eighth cohort, ages (35-39), has the highest ratio and the fifth cohort, ages (20-24), has the lowest ratio. By convention, denote this rank pattern as 85, ie., the index of the highest ratio is first and the index of the lowest ratio is second. Furthermore, let the amplitude, A, be defined as the high ratio minus the low ratio, or, in the above example,

$$A = .917 - .607 = .310$$

TABLE 4 - White Males in West Virginia
1960-1970

Age Index	Age Group	Migration Rate	Migration Ratio	Rank
1	0-4	-.033		
2	5-9	-.099		
3	10-14	-.088		
4	15-19	-.140	.860	
5	20-24	-.393	.607	Low
6	25-29	-.346	.654	
7	30-34	-.096	.904	
8	35-39	-.083	.917	High

Table 5 was set up to assign in- and out-migration patterns according to the rank pattern and the amplitude A. The pattern assignments were defined after studying plots of past migration rates for many different areas (Pittenger, 1980). The rank pattern locates the position of the peak for both the in and out migration flows - whether college(C), early(E), intermediate(I), or late(L). The amplitude (A) is used to estimate the height of the peak - short(S), tall(T), or extreme(E).

For white males in West Virginia, it was determined earlier that the rank pattern was 85 and the amplitude was (0.310). Table 5 assigns an in-migration pattern of LS and an out-migration pattern of IX. In other words, for a first guess, the in migration curve has a late peak and the peak is fairly short. The out migration curve peaks at the intermediate ages (20-24) and (25-29), and has a very high peak. Which of the three slopes should be used would depend upon the result of the calculation described in Section C.3.

TABLE 5 - Migration Pattern Assignments

Rank Pattern	Assignment
45, 46, 47, 48 (1)	CS - ET if A < .30
	CT - ET if .30 < A < .65
	CX - IX if .65 < A
56, 57, 58	ES - LS if A < .20
	ET - LT if .20 < A < .60
	EX - LT if .60 < A
64, 65, 67, 68, 54, 78	IT - LS if A < .35
	IX - LS if A > .35
74, 75, 76, 84	LS - ES if A < .20
	LT - ET if .20 < A < .60
	LX - ET if .60 < A
85, 86, 87	LS - IT if A < .30
	LS - IX if A > .30

where A = Amplitude

(1) If the ratio of the net migration ratio for age group four to that for age group eight is < 0.875, then the assignment should be that for rank pattern 86. This is to distinguish "true" college patterns from patterns more symmetrical in their outflow of young adults.

D. Retirement Migration

Before describing how to calibrate the model to obtain the desired net migration totals by race and sex, it is necessary to describe how the migration of the retirement aged population and the special populations were handled as this affects the calibration of the model.

Retirement migration is handled separately for two important reasons --

1) The migration patterns for the retirement age population can vary considerably within migration patterns exhibited by the population less than 65. The factors that cause the retirement age population to in-migrate or out-migrate are usually independent of those affecting the population less than 65.

2) Fairly good estimates of the retirement age migration can be obtained by using Medicare data on the population over 65.

D.1 Migration of the Population over 65

The 1975 county population estimates by race, sex and age (Census, 1980b) estimated the population 65 and over by using Medicare data after some adjustment factors had been applied to account for discrepancies between the counts of the 1970 Census and the 1970 Medicare data (Census, 1980b). By comparing the 1970 population survived to 1975 with the 1975 population estimates, estimates of net migration can be obtained for those cohorts over 65 (equation 11 in Appendix A). For each cohort, the net migration rate is calculated as the net migration divided by the 1970 population survived to 1975 (equation 12 in Appendix A). These net migration rates are used for the age cohorts (65-69), (70-74), and (75+). When the previously defined patterns for the age groups less than 65 are calibrated to obtain the desired net migration, these rates for the age cohorts over 65 are not altered.

D.2 Retirement Related Migration

The above paragraph describes how net migration rates were defined for the population 65 and over. However, retirement related migration does not necessarily begin at age 65. Some people, for reasons of health or finances, retire well before they reach the age of 65. Since wives are apt to be younger than their husbands, there appears to be considerable retirement related migration for females less than age 65.

Figure 4 illustrates this for areas that are well known for their in- or out-migration of the retirement age population. The upper plots show net migration rates in Arizona and Florida for males and females. The lower plots show net migration rates in New York and Illinois for males and females. In all cases, the bulge due to retirement migration starts well before the age cohort (65-69). Thus in areas experiencing large retirement migration, the migration rates of the age cohorts just below age 65 should be modified to account for this. Furthermore, it should be noted that the change in migration rates due to retirement for females preceeds that for males.

An area is considered to have "retirement" migration for a given race-sex group, if the net migration rates for all age cohorts 65 and over of that race-sex have the same sign. For example, if the net migration rate for the age cohort (65-69) is 0.25, for the age cohort (70-74) is 0.20, and for the age cohort (75+) is 0.05, this area is thought to have retirement migration. Since this retirement migration is also having some impact on the age cohorts just under age 65, the migration rates of these age cohorts must be modified accordingly. If instead, in the above example, the net migration rate for age cohort (70-74) was -0.10, the population over 65 is experiencing both in and out net migration. Since there is no strong retirement migration, no retirement related adjustment will be made to the age cohorts just below age 65.

The following modification is made for those areas that are experiencing retirement migration, whether in or out. The net migration of the population 65 and over for each race-sex group is calculated by summing over the age cohorts (65-69), (70-74), and (75+) (equation 13a in Appendix A). By using the percentages shown in Table 6, a retirement ralated migration is calculated for age cohorts (45-49) through (60-64) (equation 13 in Appendix A). These percentages were estimated from 1965-1970 Census data for selected states (Pittenger, 1980). Net migration adjustments are calculated as the net migration divided by the 1970 population survived to 1975. These adjustments are added to the estimated net migration rates defined by the model. The details of how this is done are contained in Section F.

TABLE 6 - Retirement Related Migration
as a percent of Migration of the Population over 65

Age	Males	Females
45-49	0	2.5
50-54	2.5	7.0
55-59	7.5	17.0
60-64	22.5	32.5

E. Adjustments for Special Populations

As is common with most cohort component population projection models, the special populations are handled separately (Schroeder, 1980). Since the population projections in LMPM are only concerned with the civilian population, once the military population has been subtracted out of the 1970 base population, they are left out for the rest of the projection process. The college population is also subtracted out of the 1970 base population; however, it is then added back in before obtaining the 1975 population projections. A similar process is carried out in projecting from 1975 to 1980.

However, both net migration rates and directional migration rates are usually calculated for the total resident population. The migration patterns were developed based upon age-specific migration rates that included the special populations. Census data limitations make it impossible to delete the migration of college students or the military. Thus, the patterns also include the military and student migration. After data from the 1980 Census becomes available, it may be possible to calculate estimated net migration rates for just the civilian noninstitutional population.

E.1 Military Population

The observed migration rates on which the pattern selection depends are the net migration rates from 1960 to 1970. In that time period, there were relatively few females in the military. In an area with a considerable military population, the male migration rates would be more affected by the presence of the military than would the female migration rates. It was felt that the civilian male migration rates could be better approximated by the female migration rates of the same race than by the total male migration rate of that race. Thus, in these areas, the observed male migration rates are replaced by the observed female migration rates.

E.2 Student Population

Before projecting the 1970 population, the college students are subtracted from the 1970 civilian population. The 1970 civilian noninstitutional population is then projected to 1975. Just before forcing the individual race, age and sex cohorts to sum to an independently derived population control total, the student population is added back in. Due to the lack of any more current data comparable to the 1970 Census data on students, it is assumed that the student population in a given area does not vary after 1970, ie., the student population is held fixed. The following example illustrates how this necessarily implies some student migration. The adjustment that will be made counteracts this "forced" migration. It is then up to the migration submodel to estimate the total net migration - both of students and nonstudents.

TABLE 7 - Sample of Implied Student Migration

(1) Age	(2) 1970 Pop	(3) 1970 Student Pop	(4) 1975 Survived Pop	(5) Implied Net Stud Migration	(6) Implied Net Stud Mig Rate
10-14	16,500	0	16,000		
15-19	20,000	5,000	18,000	5,000	0.2778
20-24	20,000	2,000	19,500	-3,000	-0.1539
25-29	18,000	1,000	19,400	-1,000	-0.0516
30-34	15,000	500	17,500	-500	-0.0286
35-39	14,000	0	14,500	-500	-0.0345

Since the 1975 student population is assumed to be the same as the 1970 student population, column (3), the implied net student migration, column (5), is calculated as the difference between two adjacent entries in column (3). For example, in the above, there were 5,000 students aged 15-19 in 1970 and just 2,000 students aged 20-24 in 1975. Thus, on net, 3,000 students in this cohort must have out migrated between 1970 and 1975. The net migration rate for students, column (6), is calculated as the net migration of students, column (5), divided by the 1975 survived population, column (4).

However, the observed migration rates also include the migration of students. Suppose in the above example, the observed net migration rate for the age cohort (20-24) is -0.2000. With the student migration rate of -0.1539, the migration rate for nonstudents must be -0.0461. Applying a migration rate of -0.200 along with the student submodel and hence an implied net migration of -0.1539 would lead to an overall net migration of -0.3539. To avoid this double counting of student migrants, the estimated net migration rate is adjusted by subtracting the implied student net migration (equation 10 in Appendix A). In the above example, for the age cohort (20-24), 0.1539 would be added to the estimated net migration rate. This should cancel the out migration caused by the student submodel. When the migration submodel is calibrated, it will then estimate the total net migration of the population, both students and nonstudents.

F. Calibration

Section C above described how to select an in migration pattern and an out migration pattern for each race/sex group in a particular area. Section D covered migration for the population 65 and over, and Section E covered the adjustments necessary for the special populations. This section will cover the steps involved in calibrating the patterns and in applying the various adjustments to yield the desired net migration total for this race/sex group from 1970-1975.

The following can be best understood by frequent references to Appendices A and B. The equations used are listed in Appendix A. A sample calibration for white males in the state of West Virginia is contained in Appendix B. In the following, all step and equation references are to Appendix A, any symbols used are as defined in Appendix A, and all column references are to Appendix B.

Before the in and out migration patterns are calibrated, several steps were undertaken to obtain the proper in and out migration patterns for this area.

1) By applying national survival rates to the 1970 white male population, the 1975 survived population ared 5 and over is obtained. The age cohort (0-4) is obtained by applying 1970-1975 white fertility rates for the state of West Virginia to the average number of women in each cohort of child bearing age over the five year period, and surviving these births to 1975. These numbers are printed in column (6).

2) The total 1975 population plus births occurring from 1970-1975 is the sum of column (6), or SP = 801,548.

3) From the Bureau of the Census (Census, 1980a), the estimated 1975 civilian white male population is P = 838,760.

4) By equation 1, NM = 838,760-801,548 = 37212.

5) By equation 2, NMR = 37212/801548 = 0.0464.

6) By step 3, IMR = 0.19.

7) By equation 4, OMR = 0.19-0.464 = 0.1436.

8) By step 5, the slope for the in migration curve is 0.09; the slope for the out migration curve is 0.09.

9) Table 4 illustrates step 6. The amplitude,A, is 0.310 and the rank pattern,HL, is 85.

10) By step 7, the pattern chosen for in migration is LS with a slope of 0.09 (column (2) and the pattern chosen for out migration is IX with a slope of 0.09 (column (4)).

In summary, the following steps are taken to adjust the selected in and out migration patterns so that the desired 1970-1975 migration is obtained. First, the in and out migration patterns are each scaled so that the rates for age cohort (40-44) are the percents that were estimated (equations 8 and 9). A trial net migration rate is formed by subtracting the scaled out migration pattern from the scaled in migration pattern and adding on the various adjustments for retirement and the special populations (equation 14). This trial net migration rate is applied to the 1970 population survived to 1975 plus the estimated births in that time period, to obtain an estimated net migration from 1970-1975 (equation 15). The sum over this estimated net migration is compared with the desired net migration to obtain an error term (equation 17). The scaled in migration rate is then multiplied by another scalar to account for this error. The final net migration rate is then this rescaled in migration rate minus the scaled out migration rate plus the various adjustments for retirement and the special populations.

Notice that the final adjustment was made only to the in migration rates for the population under 65. This was done on the assumption that this group is the most sensitive to changing economic conditions. Furthermore, without more data on retirement patterns and the special populations, there are no good reasons for scaling any of the adjustments made for these groups.

In more detail, the following steps are taken to calibrate the in and out migration patterns that were chosen above.

1) The in and out migration patterns are each adjusted by multiplying by a scalar (equations 8 and 9).

Basic to the pattern selection is the assumption that the in and out migration percentages for the age cohort (40-44) are the same as the in and out migration percentages for the entire race/sex group. The in migration pattern in column (2) contains 0.1750 for age cohort (40-44) whereas the estimated in migration is

0.1900. Multiplying column (2) by 0.1900/0.175 = 1.0857 gives the desired result (equation 8), which is printed in column (3). Similarly, column (4) is multiplied by 0.8206 = 0.1436/0.175 to give the estimated out migration of 0.1436 for age cohort (40-44) (equation 9). This is printed in column (5).

2) Column (7) contains the adjustment for the student migration. These numbers were obtained by applying equation 10 to the 1970 student population and to the 1975 survived population. Due to the way in which the student submodel operates (Schroeder, 1980), the entry for age cohort (15-19) is ignored by the migration submodel.

3) Columns (8) through (13) show the steps involved in obtaining the net migration rates for those 65 and over as well as the retirement related migration for age cohorts (50-54), (55-59) and (60-64). Column (8) contains the estimates of the population over 65 (Census, 1980a). By comparing these estimates with the 1975 survived population, column (6), an estimate of migration for those 65 and over, column (9), is obtained (equation 11). The estimated net migration rates, column (10), are the numbers of net migrants, column (9), divided by the survived population, column (6) (equation 12).

Since the net migration rates for the age cohorts (65-69), (70-74) and (75+) all have the same sign, it is assumed that there is some retirement related migration in the population aged (50-64). The total number of net migrants aged (65 +) can be obtained by adding the entries in column (9) and is -2215 (equation 13a). By applying the percentages given in Table 6, the retirement related migration for age cohorts (50-54), (55-59) and (60-64) can be calculated (equation 13b). The percentages from Table 6 are printed in column (11) and the results of the multiplication are in column (12). The retirement related net migration rates, column (13), are the net migrants, column (12), divided by the survived population, column (6) (equation 13c).

4) A trial net migration rate is formed as the scaled in migration rate, column (3), minus the scaled out migration rate, column (5), plus the student adjustment, column (7), plus the migration rates for the population over 65, column (10), and plus the retirement related migration, column (13) (equation 14). By applying this trial net migration rate to the

1970 population survived to 1975 plus births in the time period 1970-1975, column (6), an estimated net migration is calculated, column (15) (equation 15). At the same time, the scaled in migration rate, column (3), is applied to the 1970 population survived to 1975 plus births, column (6), to yield an estimated in migration for the population under 65, column (16) (equation 16).

5) The desired net migration total minus the net migration as just calculated (the sum of column (15)) is an error term.

$$\text{Error} = 14878 = 37212 - 22334.$$

To get rid of this error term and insure that the net migration rates yield the desired result, another scalar adjustment is made to the in migration rates for the population 65 and under, column (3). This scalar factor(F) is the total in migration of the population (the sum of column 16)) plus the error term, all divided by the total in migration of the population under 65.

$$F = (150626+14878)/150626 = 1.0988.$$

Note that if the trial net migration (the sum of column (15)) is larger than the desired, the error term is negative and the scalar factor just defined is less than one. Applying this factor to the in migration rates in column (3) would reduce them all and would thus reduce the net migration rate and the resultant net migration. Similarly, if the trial net migration is smaller than desired, the error term would be positive, the scalar factor would be greater than one, so the in migration rates and the net migration rates would increase as would the resultant net migration.

6) The final net migration rates, column (18), are calculated as the renormalized in migration rates, column (17), minus the out migration rates, column (5), plus the student adjustment, column (7) plus the migration rates for the population over 65, column (10), plus the retirement related migration, column (13) (equation 17).

7) The estimated net migration for white males, column (19), is thus the 1970 population survived to 1975 plus births that occurred from 1970-1975, column (6), times the net migration rates, column (18).

II. Implementation

The above describes how the migration model is used in projecting from 1970 to 1975. Since the population projections model in LMPM projects past 1980, the migration model is also used in projecting from 1975 to 1980. In this second stage, the necessary scaling and adjustments are done to the in and out migration patterns identified in projecting from 1970 to 1975. Note that within each area, the net migration rates as derived for each race/sex group for 1975-1980 may differ considerably from those derived for the same race/sex group for 1970-1975. In both cases the net migration rates are based upon patterns exhibited from 1960-1970. For the time period 1970-1975, the patterns are scaled and calibrated to fit an independent estimate of net migration for 1970-1975. For the time period 1975-1980, the same patterns are scaled and calibrated to an independent estimate of net migration for 1975-1980.

III. Validation

A thorough testing of this submodel can not be done until the 1980 Census data is made available. However, some measure of its performance can be obtained by comparing the population figures for 1975 with the 1975 county estimates developed by the Bureau of the Census (Census, 1980a). Although these two sets of population estimates for 1975 are based upon slightly different assumptions and are defined differently, a considerable insight can be gained by comparing them.

The population projections in LMPM are based on the 1970 Fourth Count Census data (Census, 1973) which is a count of the population based upon a 20% sample. The 1975 Census' county estimates are based upon the 1970 Second Count Census data which is a 100% count of the population. Furthermore, before projecting to 1975, these figures were updated to July 1, 1970 using July 1, 1973 total population counts. Thus the two sets do not even agree at the start. However, except for very small areas, this discrepancy is usually quite small.

A second important difference is that the LMPM projections deal with the civilian population whereas the Census' estimates deal with the resident population, ie., civilian plus military. In areas, with a large military population, eg., San Diego County, one would expect the population estimates of LMPM to be considerably smaller for the male cohorts aged (15-19), (20-24) and (25-29) than the population estimates of the Bureau of the Census for the same cohorts. However, the other cohorts that are less impacted by the presence of the military should be comparable.

Table 8 shows the results of comparing the 1975 state level population estimates of LMPM with those done by the Bureau of the Census. To compare two sets of population estimates for a particular geographic area, the mean absolute percent error was calculated by comparing the population eatimates for each race, sex and age cohort. In each comparison, the Census figures are taken as the "true" values and the LMPM figures as the "estimated" values. As was pointed out above, the two sets of estimates do not even start with the same base figures. Column (1) in Table 8 shows how the 1970 figures used by the Bureau of the Census differ from the 1970 figures used by LMPM. For most states, the mean absolute percent error is very small. It becomes larger in those states with a very small minority population. In Wyoming for example, there were 85 nonwhite males aged 70-74 in 1970 according to the Fourth Count Census, whereas there were only 59 in the adjusted Second

Count figures used by the Bureau of the Census. Although the absolute difference is quite small, the percentage difference is quite large, 44.07%.

Column (2) shows a comparison of the 1975 Census' estimates with the LMPM estimates calculated using the Plus-Minus technique, ie., before the migration submodel was developed. Column (3) shows a comparison of the Census' esimtates with the LMPM estimates calculated using the migration submodel. For each state, the entry in column (3) is less than that in column (2), and usually by a factor of 2.

TABLE 8 - Mean Absolute Percent Error

	Comparison of Base Figures for 1970	Comparison of Census '75 Estimates with '75 LMPM Estimates using	
		Plus-Minus	Migration Submodel
	(1)	(2)	(3)
Alabama	1.0	6.0	3.5
Alaska	4.7	15.5	8.2
Arizona	2.7	10.8	5.1
Arkansas	4.0	8.5	4.0
California	1.3	6.8	3.1
Colorado	2.6	10.3	4.7
Connecticut	2.9	9.8	4.1
Delaware	2.8	5.8	3.4
District of Columbia	1.9	16.1	7.9
Florida	1.6	9.8	4.8
Georgia	1.0	6.5	2.3
Hawaii	2.3	16.0	9.0
Idaho	6.0	13.3	7.5
Illinois	1.1	4.6	2.8
Indiana	1.1	4.7	2.9
Iowa	2.0	6.1	4.0
Kansas	1.4	7.2	4.3
Kentucky	1.6	6.7	4.0
Louisiana	1.1	6.1	2.5
Maine	10.6	18.7	10.4
Maryland	1.1	8.4	2.8
Massachusetts	2.8	5.8	2.8
Michigan	1.3	4.8	3.2
Minnesota	2.9	7.6	4.4
Mississippi	1.5	6.5	3.7
Missouri	1.3	5.7	3.2
Montana	4.3	9.7	5.0

TABLE 8 - Mean Absolute Percent Error (continued)

	Comparison of Base Figures for 1970	Comparison of Census '75 Estimates with '75 LMPM Estimates using	
		Plus-Minus	Migration Submodel
	(1)	(2)	(3)
Nebraska	3.5	8.3	4.0
Nevada	4.6	8.8	6.4
New Hampshire	11.6	19.5	10.9
New Jersey	1.6	7.0	3.5
New Mexico	2.9	9.2	4.8
New York	1.5	4.7	2.7
North Carolina	1.0	7.6	3.1
North Dakota	6.2	12.7	8.6
Ohio	1.1	4.2	3.0
Oklahoma	1.5	7.7	3.3
Oregon	2.4	8.0	4.2
Pennsylvania	1.1	5.1	2.7
Rhode Island	4.9	9.4	5.6
South Carolina	1.4	9.3	3.6
South Dakota	3.5	9.0	6.6
Tennessee	1.1	5.8	6.3
Texas	1.0	6.8	2.3
Utah	4.0	11.3	6.3
Vermont	14.1	20.1	15.2
Virginia	1.0	8.0	3.5
Washington	2.8	7.7	4.3
West Virginia	2.6	9.7	3.8
Wisconsin	1.8	12.4	3.7
Wyoming	8.2	14.7	8.1

IV. Future Improvements

A considerable amount of work remains to be done. The above represents only a first attempt at implementing a computerized, flexible migration model. Several aspects of this model warrant further study; in particular, some of the more important are -

1) The 36 patterns that are listed in Table 1 and which the model uses to generate current migration flows could probably be refined.

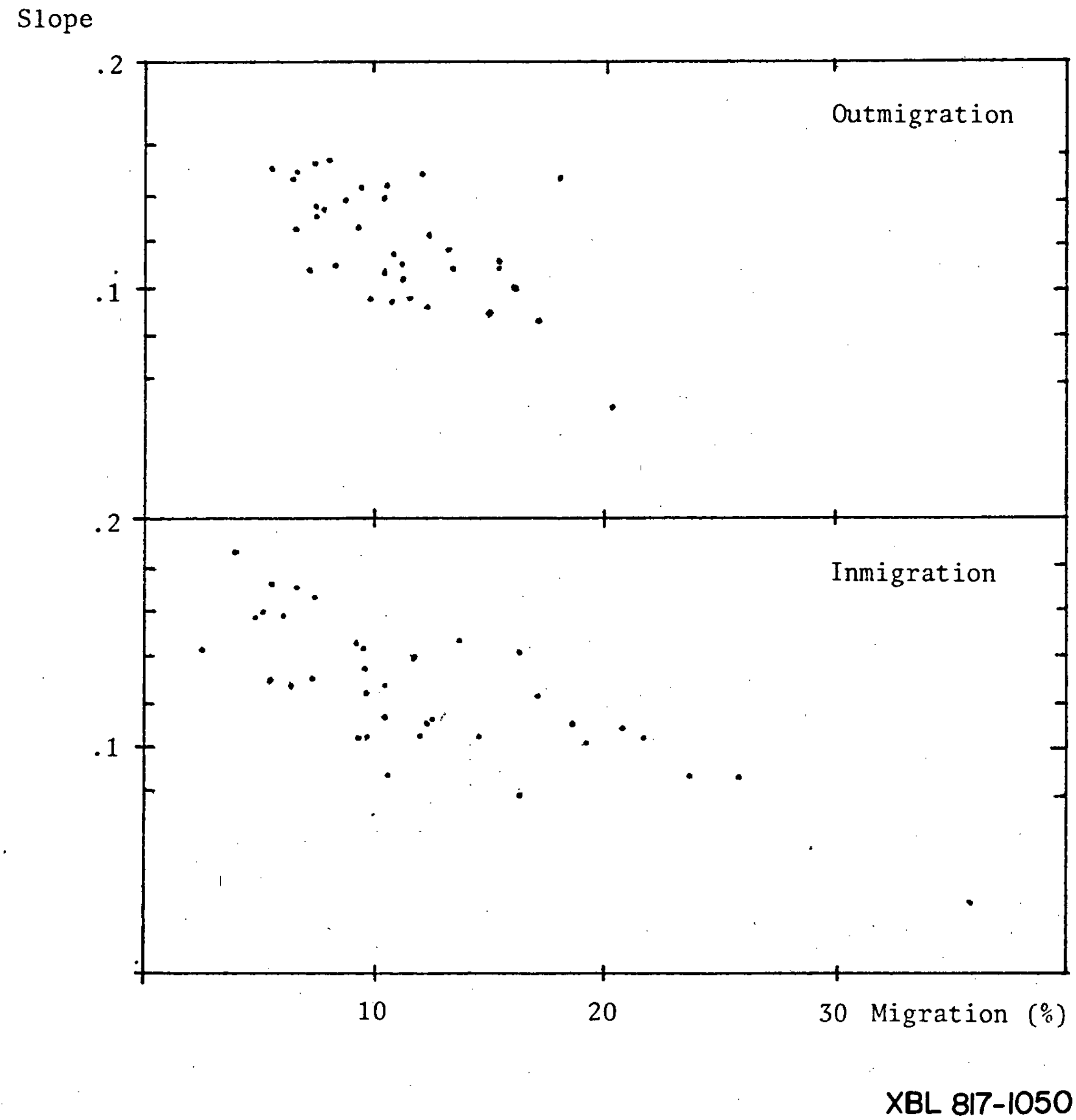
2) The way in which the in and out migration patterns are chosen may be improved upon.

However, most of this work awaits the release of the 1980 Census. Not until the population projections produced with this migration submodel are compared with the population counts from the 1980 Census, can the real strengths and weaknesses of this model be pointed out. At that time, it should be possible to improve upon the definitions of the patterns currently used and maybe also the way in which the patterns are chosen.

ACKNOWLEDGEMENT

This work was supported by the U.S. Department of Energy under Contract No. W-7405-ENG-48.

Figure 1.--Relationship of Migration Slope to Migration Level at Ages 40-44: Selected State Economic Areas, 1965-70; Males.



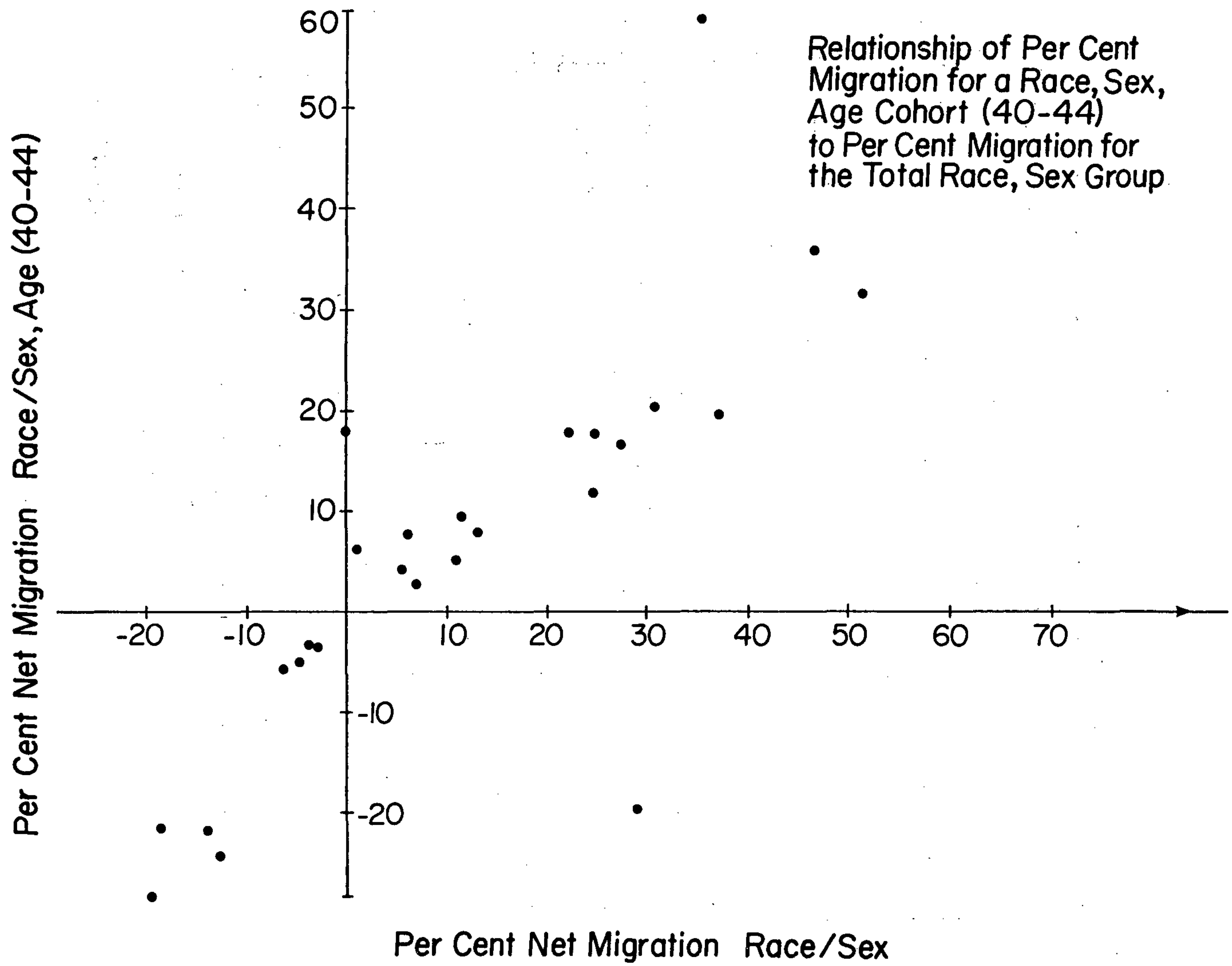
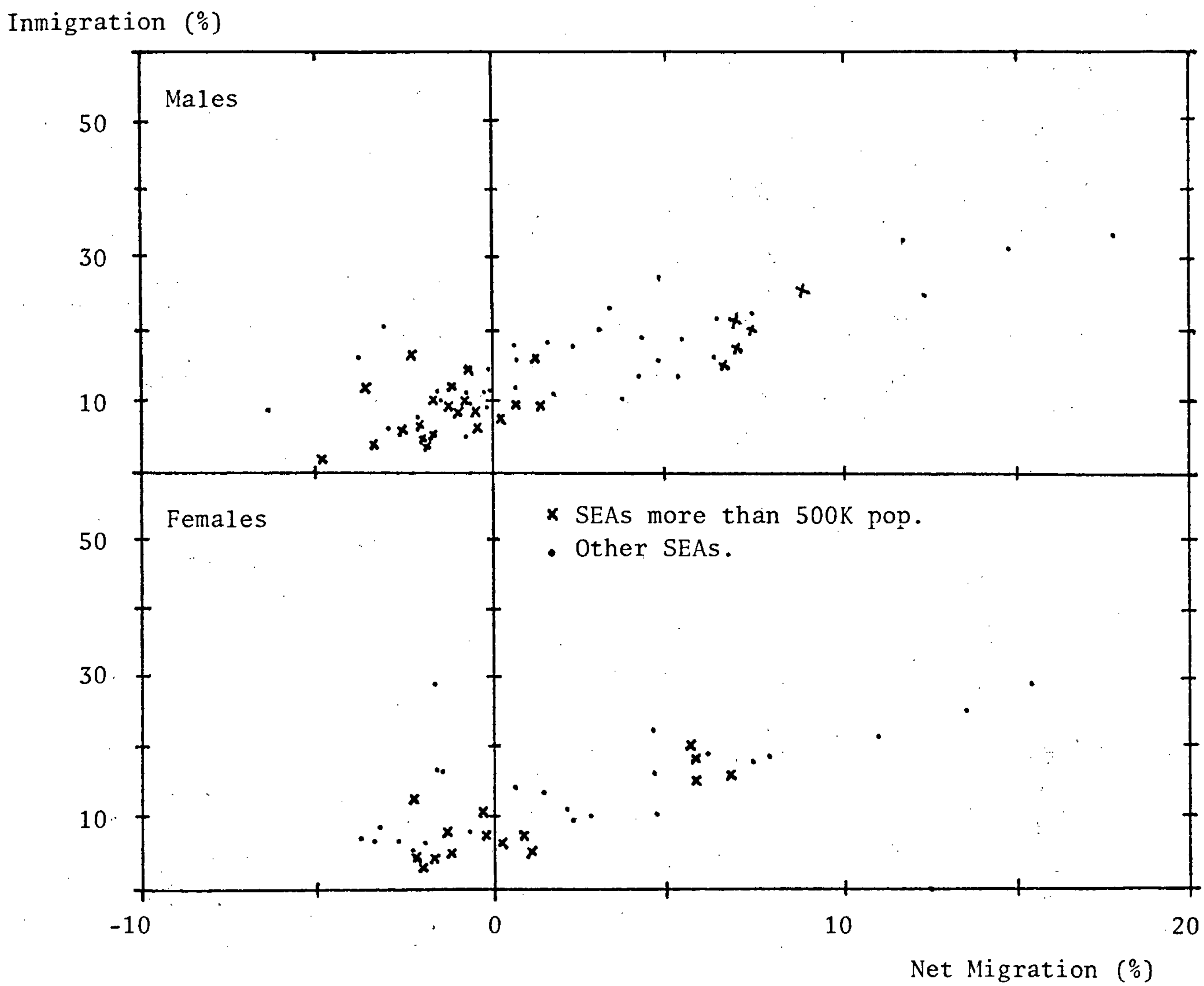


Figure 2

XBL 817-1051

Figure 3.--Relationship of Inmigration to Net Migration, Ages 40-44: Selected State Economic Areas, 1965-70.



XBL 817-1052

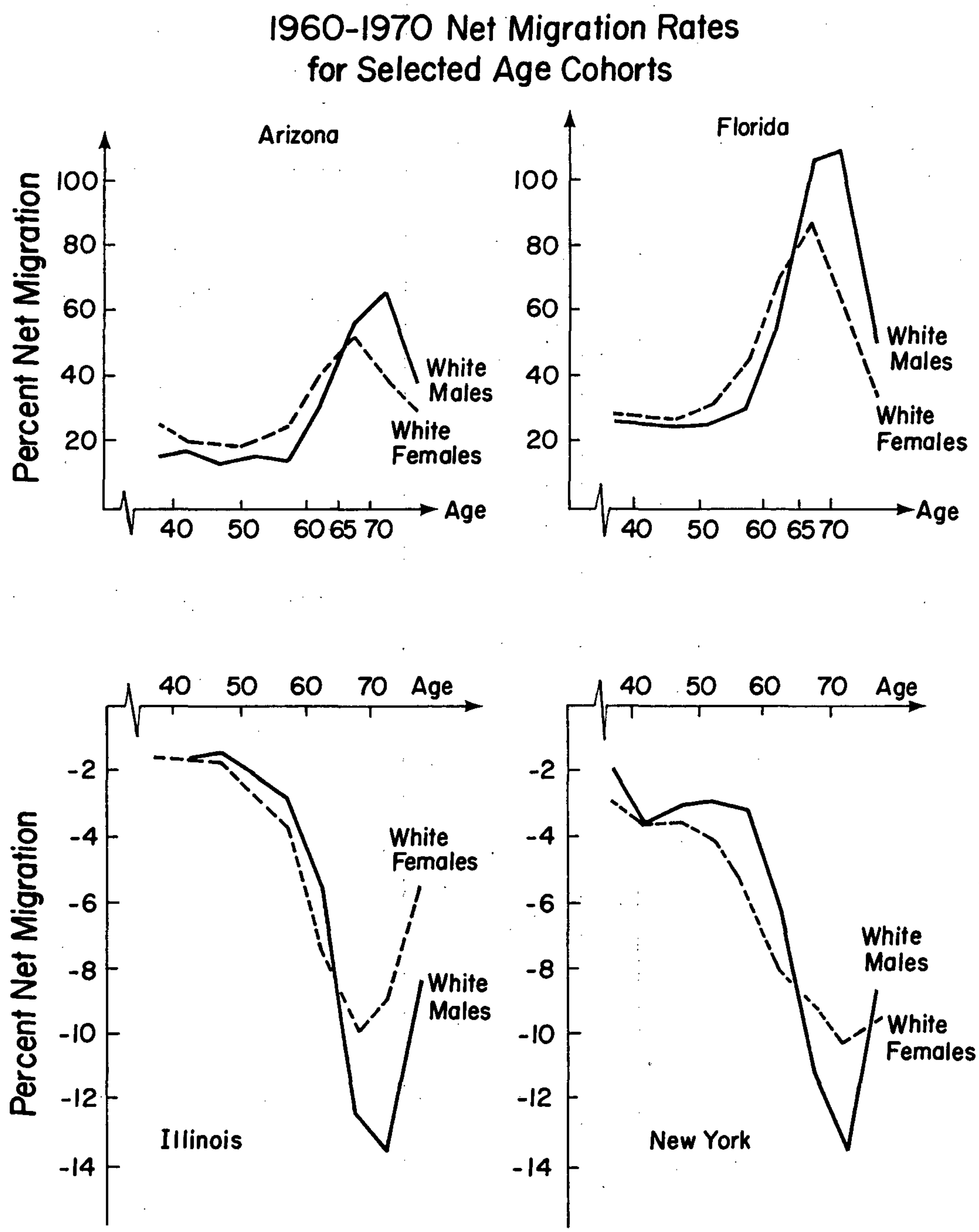


Figure 4

XBL 817-1053

VI. References

Bowles, Gladys, Calvin L. Beale, and Everett S. Lee, 1975, "Net Migration of the Population, 1960-1970, by Age, Sex, and Color, United States, Regions, Division, States, and Counties", Population-Migration Report 1960-1970, Parts 1-6, Economic Research Service, U. S. Department of Agriculture; The Institute for Behavioral Research, University of Georgia; and Research Applies to National Needs, National Science Foundation, cooperating

Pittenger, Donald B., 1974, "A Typology of Age-Specific Net Migration Rate Distributions", Journal of the American Institute of Planners, July 1974, Vol. 40, No. 4

----------, 1978, "On Making Flexible Projections of Age-Specific Net Migration", Environment and Planning A, 1978, Vol. 10, pp 1253-1272

----------, fall 1980, personal correspondence

Schroeder, Esther, August 1980, "The Labor Market Projections Model - A User's Guide to the Population, Labor Force, and Unemployment Projections Model at Lawrence Berkeley Laboratory", LBL-11349

Shryock, Henry S., Jacob S. Siegel, and Associates, 1973, The Methods and Materials of Demography, U. S. Bureau of the Census, U. S. Government Printing Office

U. S. Bureau of the Census, 1963, U. S. Census of Population 1960. Subject Reports. "Mobility for States and State Economic Areas", Final Report PC(2)-2B. Washington, D.C., U. S. Government Printing Office

----------, 1972, U. S. Census of Population 1970. "Migration Between State Economic Areas", Final Report PC(2)-2E. Washington, D.C., U. S. Government Printing Office

----------, 1973, 1970 Fourth Count (Population) Summary Tapes

----------, 1976, Revenue Sharing tape obtained from the Customer Services Branch, Bureau of the Census, Washington, D.C.

----------, 1980a, County Population Estimates by Age, Sex, and Race for 1975, Estimates Research Unit, Population Division, Washington, D.C.

VI. References

----------, 1980b, Current Population Report, Series P-23, No. 103, "Methodology for Experimental Estimates of the Population of Counties, by Age and Sex: July 1, 1975", U.S. Government Printing Office, Washington, D.C.

Appendix A

Equations

used by

the Migration Submodel

Appendix A

Notation-

To simplify the notation, no indeces are included for race or sex. Instead, the following is meant to apply to a particular race/sex group in a given area.

Symbol	Definition
i	index of five year age cohort, i=1,2,...,16
SP	1970 population survived to 1975, by applying a survival rate to the 1970 population and to births occurring from 1970-1975
ST_i	student population in 1970 (also in 1975) for i=4,...,7
P	1975 population
NM	1970-1975 net migration
NMR	1970-1975 net migration rate
IM	1970-1975 in migration of the population <65, and not as a result of any adjustments
IMR	1970-1975 in migration rate
ØMR	1970-1975 out migration rate
SI	slope of in migration pattern
SØ	slope of out migration pattern
$M67_i$	1960-1970 net migration rate, i=1,...,16
IP_i	in migration pattern, i=1,...,13
$ØP_i$	out migration pattern, i=1,...,13
AIP_i	adjusted in migration pattern, i=1,...,13
$AØP_i$	adjusted out migration pattern, i=1,...,13
SA_i	student adjustment, i=5,...,8
RM_i	retirement migration i=14,...,16
RMR_i	retirement migration rate for i=14,...,16 related retirement migration rate for i=10,13
RRF_i	retirement related migration adjustment factor for i=10,13
Λ	an estimate

Equations-

Please note-

a) In the following, any symbol without a subscript stands for a total; any symbol with a subscript (i is the only subscript), stands for a number or rate for the ith age cohort.

b) Some of the rates are defined for only some age cohorts. For simplification of notation, these rates should be considered as zero for those cohorts for which they are not defined.

---

1) NM = P-SP

2) NMR = NM/SP

3) Determine IMR from NMR and Table 2

4) OMR = IMR-NMR

5) Determine SI from IMR and Table 3
Determine SØ from ØMR and Table 3

6) Calculate $1 + M67_i$ for $i=4,8$

Let $H = i \ni 1 + M67_i$ is a maximum
Let $L = i \ni 1 + M67_i$ is a minimum

Then $A = M67_H - M67_L$
and rank pattern is HL

7) Select IP_i for $i=1,\ldots,13$ from HL, A, Table 5, SI and Table 1
Select $ØP_i$ for $i=1,\ldots,13$ from HL, A, Table 5, SØ and Table 1

8) $AIP_i = (SI/IP_9) \times IP_i$ for $i=1,\ldots,13$

9) $AØP_i = (SØ/ØP_9) \times ØP_i$ for $i=1,\ldots,13$

10) $SA_i = -(ST_i - ST_{i-1})/SP_i$ for $i=5,\ldots,8$

11) $RM_i = P_i - SP_i$ for $i=14,\ldots,16$

12) $RMR_i = (P_i - SP_i)/SP_i$ for $i=14,\ldots,16$

13) If RMR_i for i=14,...,16 are both plus and minus, set $RMR_i = 0$ for i=10,...,13 and skip to step 14)

If RMR_i for i=14,...,16 all have the same sign, then

13a) $RM = \sum_{i=14}^{16} RM_i$

13b) $RM_i = RRF_i \times RM$ for i=10,...,13

13c) $RMR_i = RM_i/SP_i$ for i=10,...,13

14) $\hat{NMR}_i = AIP_i - A\emptyset P_i + SA_i + RMR_i$ for i=1,...,16

15) $\hat{NM} = \sum_{i=1}^{16} NMR_i \times SP_i$

16) $\hat{IM} = \sum_{i=1}^{13} AIP_i \times SP_i$

17) Let $Err = NM - \hat{NM}$

Then $NMR_i = \left(\frac{\hat{IM} + Err}{\hat{IM}}\right) \times AIP_i - A\emptyset P_i + SA_i + RMR_i$ for i=1,2,...,16

Appendix B

A Calibration

of the Migration Submodel

for the time interval 1970-1975

for White Males

in West Virginia

	IP_i	AIP_i	$ØP_i$	$AØP_i$	SP_i	SA_i	P_i	RM_i	RMR_i	RRF_i	RM_i	RMR_i	$\hat{NMR}_i$	$\hat{NM}_i$	$\hat{IM}_i$		NMR_i	NM_i
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
(0-4)	.1490	.1618	.1490	.1222	59798								.0396	2368	9675	.1778	.0556	3325
(5-9)	.2450	.2660	.2450	.2010	67562								.0650	4392	17971	.2923	.0913	6168
(10-14)	.1980	.2150	.1980	.1625	78136								.0525	4102	16799	.2362	.0737	5759
(15-19)	.1800	.1954	.2200	.1805	85595								.0149	1275	16725	.2147	.0342	2927
(20-24)	.2800	.3040	.5800	.4759	72932	.0759							-.0960	-7001	22171	.3340	-.0660	-4814
(25-29)	.3260	.3539	.5200	.4266	43825	.2650							.1923	8428	15510	.3889	.2273	9961
(30-34)	.2650	.2877	.2900	.2379	46080								.0498	2295	13257	.3161	.0782	3603
(35-39)	.2150	.2334	.2150	.1764	41767								.0570	2381	9748	.2565	.0801	3346
(40-44)	.1750	.1900	.1750	.1436	40048								.0464	1858	7609	.2088	.0652	2611
(45-49)	.1420	.1542	.1420	.1165	47071								.0377	1775	7258	.1694	.0529	2490
(50-54)	.1160	.1259	.1160	.0952	49047					.025	-55	-.0011	.0296	1452	6175	.1383	.0420	2060
(55-59)	.0940	.1021	.0940	.0771	43481					.075	-166	-.0038	.0212	922	4439	.1122	.0313	1361
(60-64)	.0764	.0829	.0764	.0627	39679					.225	-498	-.0126	.0076	302	3289	.0911	.0158	627
(65-69)					32932		32436	-496	-.0151				-.0151	-496			-.0151	-496
(70-74)					22741		22196	-545	-.0240				-.0240	-545			-.0240	-545
(75+)					30854		29680	-1174	-.0381				-.0381	-1174			-.0381	-1174
Total					801548			-2215						22334	150626			37209

(1) Age at end of projection interval
(2) In-migration pattern LS with slope .09
(3) Col (2) times 1.0857 = 0.19/0.175, equation 8
(4) Out-migration pattern IX with slope .09
(5) Col (4) times 0.8206 = 0.1436/0.175, equation 9
(6) 1975 Survived Population and survived births
(7) Adjustment for student migration
(8) Estimated 1975 population for age cohorts (65-69) and older
(9) Estimated migrants 65 and over, Col(8) - Col(6), equation 11
(10) Estimated net migration rates for the population 65 and over, Col(9)/Col(6), equation 12
(11) Percents to estimate retirement population, from Table 6
(12) Retirement related migration = Col(11) x (-2215) where -2215 = sum of Col(9), equation 13b
(13) Retirement related net migration rates = Col(12)/Col(6), equation 13c
(14) Trial Net Migration Rate = Col(3) - Col(5) + Col(7) + Col(10) + Col(13), equation 14
(15) Trial Net Migration = Col(6) x Col(14), equation 15
(16) In Migration = Col(3) x Col(6), equation 16
(17) Renormalized in Migration Rate = Col(3) x 1.0988
where 1.0988 = (Sum Col(16) + (Desired Mig - Sum Col(15)))/Sum Col(16)
(18) Net Migration Rate = Col(17) - Col(5) + Col(7) + Col(10) + Col(13), equation 17
(19) Net Migration = Col(6) x Col(18)

TECHNICAL INFORMATION DEPARTMENT
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720