LBL-13587
Preprint

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

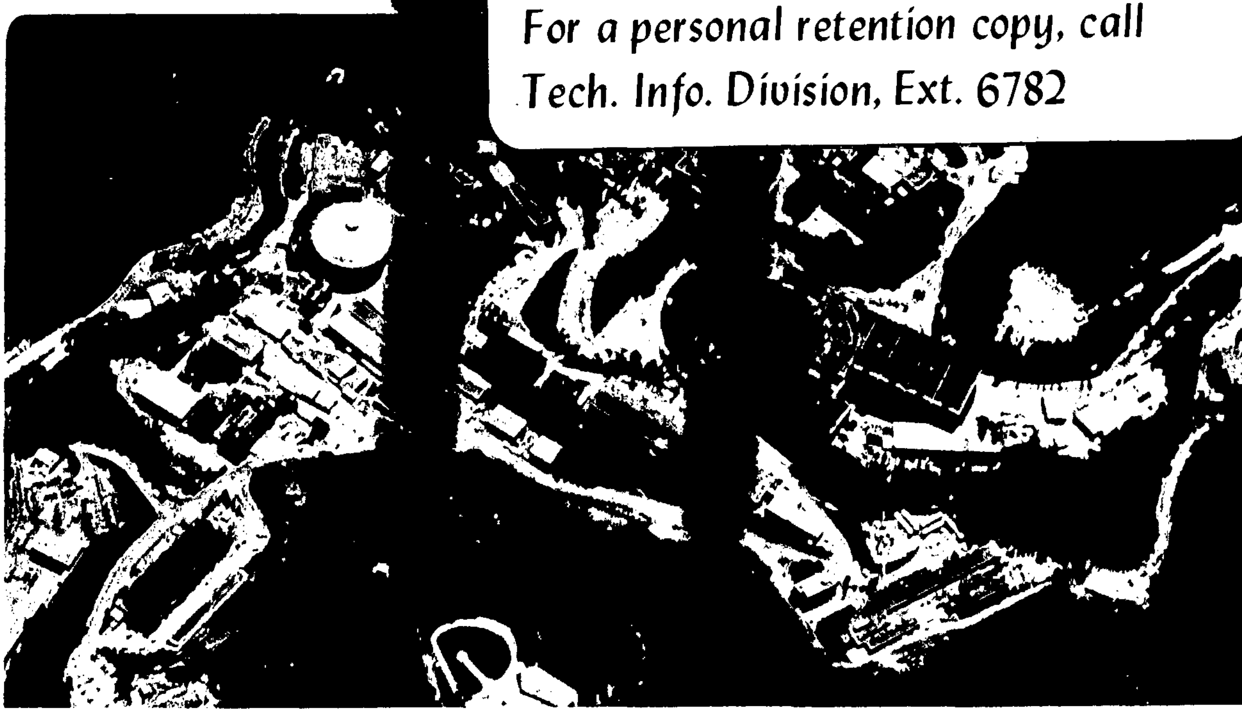
Physics, Computer Science & Mathematics Division

Submitted for publication

A TECHNIQUE FOR IMPROVING THE ACCURACY OF MIGRATION AGE DETAIL IN MULTIPLE-AREA POPULATION FORECASTS

Esther C. Schroeder and Donald B. Pittenger

November 1981



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

DISCLAIMER

LBL-13567

A Technique for Improving the Accuracy

of Migration Age Detail

in Multiple-Area Population Forecasts

November 1981

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

Donald B. Pittenger
The Demographics Laboratory
2065 Lakemoor Dr., S.W.
Olympia, Wa., 98502

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

A Technique for Improving the Accuracy

of Migration Age Detail

in Multiple-Area Population Forecasts

Abstract

Population projections are often required for many geographical areas, and must be prepared with maximal computer and minimal analytical effort. At the same time, realistic age detail forecasts require a flexible means of treating age-specific net migration. This report presents a migration projection technique compatible with these constraints. A simplified version of Pittenger's model is used, where future migration patterns are automatically assigned from characteristics of historical patterns. A comparative 1970-1980 test of age pattern accuracy indicates that this technique is superior to the commonly used plus-minus adjustment to historical rates.

Table of Contents

I. Introduction

There is no universally superior method of treating migration in population projections. This assertion is made from an <u>applied</u> rather than <u>ideal</u> standpoint. If such constraining factors as time, cost, computer storage and data manipulation capabilities were disregarded, it is relatively easy to imagine an ideal migration component. But constraints exist and demographers must deal with them.

This paper addresses the situation where age-specific net migration rates must be determined for many geographic areas with a minimum expenditure of analytical effort. The requirement of high production using small staff is common in the public, private, and academic sectors, and its constraining elements are obvious. Added to this is the desire to treat age-specific migration flexibly; flexible migration rates should yield more accurate age detail than would relatively fixed rates based on historical migration patterns. However, flexibility is usually accompanied by model complexity, whereas high production and few personnel would seem to dictate simplicity.

The subject of this paper is one solution to the problem just posed. We shall present a technique for projecting race-sex-age-specific migration that is more sophisticated than most techniques presently in use, yet has the capability of being used in mass-production situations. Although the solution has some limitations, it seems to offer a significant advance in the craft of population projection.

A. Existing Techniques for Treating Migration

The problem indicated in the last section included a requirement for race-sex-age detail in population projection output. This means that we are dealing with techniques used under the "cohort-component" methodological framework. This methodology refers to the case where populations are broken down into race-sex-age groups and moved through time by multiplication by race-sex-age specific rates of fertility, mortality and migration. For details, consult Irwin (1977), Pittenger (1976) or Shryock, Siegal and Associates (1973).

Until recently, most demographers or other technicians preparing sub-national projections have relied on the assumption that future race-sex-age-specific net migration rate patterns will be similar to historical patterns for the population in question (Pittenger, 1976, Chapter 8). Irwin's (1977) Census Bureau manual for local planners provides an example of a technique for modifying historical decade migration rates using a "plus-minus adjustment" so that known post-censal trends may be accomodated (also see Shryock et al., pp 705-6).

The Plus-Minus Technique stretches or compresses net migration rates for individual age groups so that a desired net migration total can be obtained when these rates are applied to the population. A plus factor is applied proportionately to all positive rates that increases them if more in-migration is desired, or decreases them if less in-migration is desired. A minus factor acts in a similar manner on the the negative rates.

This 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, and often unrealistically.

The need for better means of treating age-specific net migration has led to improved models in recent years. The Bureau of the Census (1979) has, since the 1960s, used a model that projects state out-migration by race, age and sex into a pool for reallocation as in-migration. While conceptually attractive, it is operationally of limited flexibility because the flow patterns are fixed in form. Also, the model cannot easily be applied to areas smaller than metropolitan areas and State Economic Areas because historical county migration flow data in five-year age detail have not been released by the Census Bureau.

Andrei Rogers (1975) has proposed highly complex multi-regional accounting models that deal with migration in terms of age-specific flows from one region to another. This is attractive as the analyst is dealing with true "at risk" rates, and not pseudo-rates such as rates of in or net migration. Unfortunately, the models designed by Rogers are so complex that data cannot always be found to make them operational. Related to this is an apparent lack of flexibility. Most analysts faced with the task of generating numbers will have to make do with other techniques until Rogers and his associates come up with simpler, more practical methods.

I. Introduction

Pittenger (1978) proposed a technique for flexibly modeling age-specific net migration rate patterns. It is based on a simple typology of underlying directional (in and out) age-specific migration rate patterns. Although intermediate cases are possible, directional flow patterns tend to have either early or late age timing of the modal rates. On a five year model, "early" might be age group 20-24 and "late" could be age group 25-29 or 30-34. One class of net patterns ("younger") occurs when in-migration timing is early and out-migration has a late modal age. Another case ("older") has a late in-migration mode and an early out-migration mode. Shapes of net rate patterns within each type will vary depending upon the magnitude of the rates in each direcion. For example, high age-specific out-migration rates combined with low in-migration rates yield net patterns that are out-migratory. These concepts are illustrated in Figure 1.

Pittenger has also observed that the directional flow timing patterns for areas as small as counties normally do not change over time, even though overall net migration moves from highly out to highly in. This means that, once timing patterns have been established for a given population, realistic age-specific net migration rate patterns can be modeled simply by increasing or decreasing rate magnitudes for directional flows. His 1978 paper describes a model in which the migration rates automatically adjust to changes in a single, exogeneously specified variable such as total labor force, total net migration, or occupied housing stock.

Pittenger's model is conceptually elegant in its simple treatment of complicated observed net migration patterns. It has two weaknesses however. First, unlike the Census Bureau and Rogers models, migration for individual populations is not controlled to national totals. Thus the sum of projected race-sex-age cohorts over substate areas would probably not equal an independently projected corresponding race-sex-age cohort for the state. This weakness can be minimized by forcing the resultant population projections for each race-sex-age cohort to add up to the corresponding race-sex-age cohort for the next larger geographic unit.

The second weakness is that the computation algorithm presented in the 1978 paper is complicated and requires a great deal of analytical effort in assigning parameter values and calibration testing. When projecting a few populations, this creates no special difficulty. But it presents a significant problem if the projections are to be mass produced, which is the concern of this paper.

B. Methodological Strategy

From a practical standpoint, the Pittenger model seemed to be the best of the flexible techniques because its data requirements appeared to be simpler. However, there were problems to be solved before it could be used to mass produce population projections with limited personnel.

A major problem entailed simplifying the computation algorithm. Details of this solution are presented below; however, this simplification was acheived at the cost of reducing the model's flexibility. The flexibility loss is fairly small, as it is confined to limiting the detail of the migration rate patterns around the ages where inflection points occur.

The most difficult problem had to do with the assignment of migration pattern types to individual populations. An analyst working with a single population can perform this task in a few minutes. If, however, population projections are being calculated for hundreds of different geographic areas, the analytical time expenditure becomes prohibitive. Therefore, an automated assignment procedure had to be developed to replace an essentially judgmental task.

II. Model Implementation

A. Defining Migration Patterns

Due to the large number of areas which can be handled by this model, it is often impossible to study each individually. However, by analyzing certain key data items, it is possible to approximate directional age-specific migration rates.

To estimate the age rate pattern for the net migration total for a race-sex group, the following procedure is used. The model determines the migration 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 defined only up to age 65.

The following discussion will concentrate on how this phase of the model was implemented. The technique established migration "character" using historical age-specific net migration rates as the reference. Specifically, it operates from intercensal rates calculated using common census-survival techniques. This means the model can be calibrated for geographical units ranging in size from states to metropolitan areas, counties, cities, and even down to census tracts. Published data for 1960-1970 are available for the first three types of units (Bowles et al., 1975); rates for the remaining units can be calculated readily.

The key to the understanding and modeling of age-specific net migration rates 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 thus 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.

Model migration patterns were designed with reference to three factors - the height of the peak age-specific rate, the age at which this peak occurs, and the general height of the curve as defined by the rate for age group 40-44. Age group 40-44 is used because its migratory behavior is comparatively unaffected by college attendance, military service and retirement. The goal was to be able to mimic most migration curves with a limited number of fixed pattern shapes, rather than trying to use a parameterized model to calculate the shapes as in Pittenger (1978). Thirty-six patterns were created for the present application. For each of three height categories, there are three amplitudes at the peak (short, tall and extreme), and four timings of the peak ("college", "early", "intermediate" and "late").

The source for analyzing age-specific directional migration flows was data published for 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 2 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 out-migration and the lower, in-migration. Both show that at ages 40-44, the slope of the directional migration rate becomes less steep with increasing migration. Net migration rates for this age group empirically are often similar to net rates for all age groups combined. To preserve the relationship of slope to migration rate level, 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.

Since most of the migration occurs in the ages 15-34, the migration model is most concerned with estimating the migration flows there. At present, it does not seem necessary to distinguish different migration patterns within each slope value for the remaining age groups. Thus, in Table 1, the number that is printed under CS for these age groups is identical across all patterns within the slope category. Further study may indicate that there should be different definitions for age groups less than 15 or for those age groups 35 and over.

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

	Pattern Type											
Ages	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

B. Pattern Assignment

The preceeding section defined the migration patterns that are used by the migration model. This section explains how particular in- and out-migration patterns are chosen for each race-sex group in a given area. Since migration flows in any area are apt to be considerably different for each such group, the following pattern assignment procedure is used once for each group. 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 follow if the reader also refers to this Appendix.

B.1 Estimating Overall Net Migration

The first step is to determine the total net migration for each race-sex group for the initial projection or estimation interval. Since our projections will later be tested for accuracy over the 1970-1980 decade, we need an exogenous forecast of the 1975 population by race-sex. The volume of migration 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. The ratio of this total net migration to the 1970 population survived to 1975 is the total net migration rate.

B.2 Estimating In and Out Migration Rates

In defining the directional migration flow patterns, considerable attention was paid to the rates for age group 40-44 (Pittenger, 1978). To decide which pattern is applicable it is necessary to determine directional migration rates and slopes at this age group. Figure 3 shows a plot of net migration rates for age group 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 group 40-44 can be approximated by that of the total race-sex group, i.e., the percent net migration at age 40-44 is very close to the percent net migration of the entire race-sex group in many areas.

It is now necessary to go from the percent net migration at age 40-44 to the percent in-migration at age 40-44 and the percent out-migration at age 40-44. Figure 4 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 - Assignment of In-Migration Rates for Ages (40-44) Given Net Migration for Ages (40-44)

Net Migration Rate	In-Migration Rate Male	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%

Table 2 is a tabular form of the data plotted in Figure 4 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.

B.3 Estimating the Slope

As was mentioned in Section A 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 is a tabular form of the data plotted in Figure 2.

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

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.

B.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 as to 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 groups 15-19 through 35-39. Since these are the age groups 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 here. Let the net migration ratio be defined as one plus the net migration rate. Then we can calculate the net migration ratio for age cohorts 15-19 through 35-39, i.e., for age groups 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).

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

Of the five age groups of interest, the eighth group, ages 35-39, has the highest ratio and the fifth group, 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 this example,

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

Table 5 was established 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. 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).

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, for rank patterns 45, 46 or 47, the ratio of the net migration ratio for age group 4 to that for age group 8 is <1.125, 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.

For white males in West Virginia, it was determined 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 that is fairly short. The out-migration curve peaks at the intermediate ages 20-24 and 25-29 and is very high. The slope assignment would depend on the result of the calculation described in Section B.3.

C. Retirement Migration

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 retirement age migration can be obtained by using Medicare data on the population over 65, if the model is to be used for making intercensal or postcensal age estimates rather than for forecasting.

C.1 Migration of the Population over 65

The 1975 county population estimates by race, sex and age treated the population 65 and over by using Medicare data after 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 that survived to 1975 with the 1975 Medicare based estimates, estimates of net migration can be obtained for those cohorts over 65. For each cohort, the net migration rate is calculated as the net migration divided by the 1970 population survived to 1975. These net migration rates are used for the age groups 65-69, 70-74, and 75+.

C.2 Retirement Related Migration

The above paragraph describes how net migration rates are 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 finance, 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 5 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 group 65-69. Thus in areas experiencing large retirement migration, the migration rates of the age groups 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 groups 65 and over of that race-sex have the same sign. Since this retirement migration is also having some impact on the age groups just under age 65, the migration rates of these age groups must be modified accordingly. If instead, the net migration rates for the age groups over 64 are both positive and negative, there is no strong retirement migration and therefore no retirement related adjustment will be made to the age groups 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 groups 65-69, 70-74, and 75+. By using the percentages shown in Table 6, a retirement related migration is calculated for age groups 45-49 through 60-64.

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

These percentages were estimated from 1965-1970 Census data for selected states. 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 E.

D. 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 the project that funded this research 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 based on census data are usually calculated for the total resident population. Census data limitations make it impossible to delete the migration of college students or the military. Thus, the 1960-1970 patterns from which the model is calibrated in this example, also include the military and student migration. After detailed data from the 1980 Census become available, it may be possible to calculate estimated net migration rates for just the civilian non-institutional population.

D.1 Military Population

The observed migration rates on which the pattern selection depends, in this illustration, are the net migration rates from 1960 to 1970. In that period, there were relatively few females in the military. In an area with a considerable military population, male migration rates would be more affected by the presence of the military than would 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 rather 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.

D.2 Student Population

Before projecting the 1970 population, the college students are subtracted from the 1970 civilian population. The 1970 civilian non-institutional population is then projected to 1975. Just before forcing the individual race-age-sex cohorts to sum to an independently derived population control total, the student population is added back in. Due to the lack of 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, i.e., the student population is held fixed. Users of the model may choose to incorporate data based on an alternative assumption. Table 7 illustrates how this process necessarily implies some student migration. The adjustment that will be made counteracts this "forced" migration. It is then up to the migration model to estimate the total net migration -- both of students and non-students.

TABLE 7 - Sample of Implied Student Migration

(1) Age	(2) 1970 Pop	(3) 1970 Student Pop	(4) 1975 Survived Pop	(5) Net Implied Student Migration	(6) Net Implied Student Mig Rate
10-14	18,500	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, or column (3), the implied net student migration, or column (5), is calculated as the difference between two adjacent entries in column (3). 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 group 20-24 is -0.2000. With the student migration rate of -0.1539, the migration rate for non-students must be -0.0461. Applying the migration model along with the student model yields a total migration rate of -0.3539 (-0.2000 from the migration model and -0.1539 from the student submodel). To avoid this double counting of student migrants, the estimated net migration rate is adjusted by subtracting the implied student net migration. In the above example, for the age group 20-24, 0.1539 would be added to the estimated net migration rate. This should cancel the out-migration caused by the student model. When the migration model is calibrated, it will then estimate the total net migration of the population, both students and non-students.

E. Calibration

This section covers the steps involved in calibrating the patterns and applying the various adjustments to yield the desired net migration total. 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. By referring to these appendices, the steps taken to calibrate the model should become clear.

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 group 40-44 are the percent in- and out-migration, respectively, that were estimated as described in Section B.2 above. A trial net migration rate vector 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. This trial net migration rate vector 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. The sum over this estimated net migration is compared with the desired net migration to obtain an error term. The scaled in-migration rates are then multiplied by another scalar to correct for this error. Each final net migration rate is the 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 is made only to the in-migration rates for the population under 65. This was done on the assumption that this group is more sensitive to changing economic conditions than the retired population. 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.

The above describes how the migration model is used in projecting from 1970 to 1975. When projecting from 1975 to 1980, 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 period 1970-1975, the patterns are scaled and calibrated to fit an independent estimate of net migration for 1970-1975. For the period 1975-1980, the same patterns are scaled and calibrated to an independent estimate of net migration for 1975-1980. Beyond 1980, net migration rate targets would be exogenously forecast. When requisite 1980 census data are available, 1970-1980 age-specific net migration rates would be used to calibrate the model, and 1980 would be the starting benchmark.

III. Validation

A thorough testing of this model can not be done until the complete 1980 Census is available. However, some measure of its performance can be obtained by comparing the 1980 state population projections with the available 1980 figures on state population by age (Census, 1981). Our computerized population projections model was run twice for each state -- once using the migration model just described and once using the plus-minus technique. The plus-minus technique was applied to the observed 1960-1970 age-specific net migration rates (Bowles et al., 1975), after these rates had been divided by two to obtain half decade rates. This procedure is not uncommon.

Figure 6 shows the results of comparing both of our 1980 state level population estimates with the early results from the 1980 Census. The first two columns indicate the mean absolute percent error over all age groups. The two remaining columns present the percent error for the age group having the greatest relative difference between the projection and the census result.

The mean absolute percent error of the migration submodel is less than that of the plus-minus technique in all but two states. Furthermore, it is usually less by almost a factor of two. Looking at the maximum error, again the migration model outperforms the plus-minus technique and usually by a factor of two. For several states, Illinois, Louisiana, Montana, New Hampshire, New Jersey and Texas, the percent error of the worst fitted age group using the migration model is <u>less than</u> the mean absolute percent error using the plus-minus technique!

In calculating the 1980 population figures, the projections were controlled to the 1980 census totals for each state in order to isolate the effects of the migration techniques with respect to age detail. As was mentionned above, military populations were deleted from the data for 1970 and were not included in the 1980 projections in either technique. The 1980 census figures do include military personnel which could not be subtracted out. However, it was felt that the errors due to this incomparability were quite small in most states. A more complete evaluation of the migration model cannot be made until more of the 1980 census data are available.

IV. Future Improvements

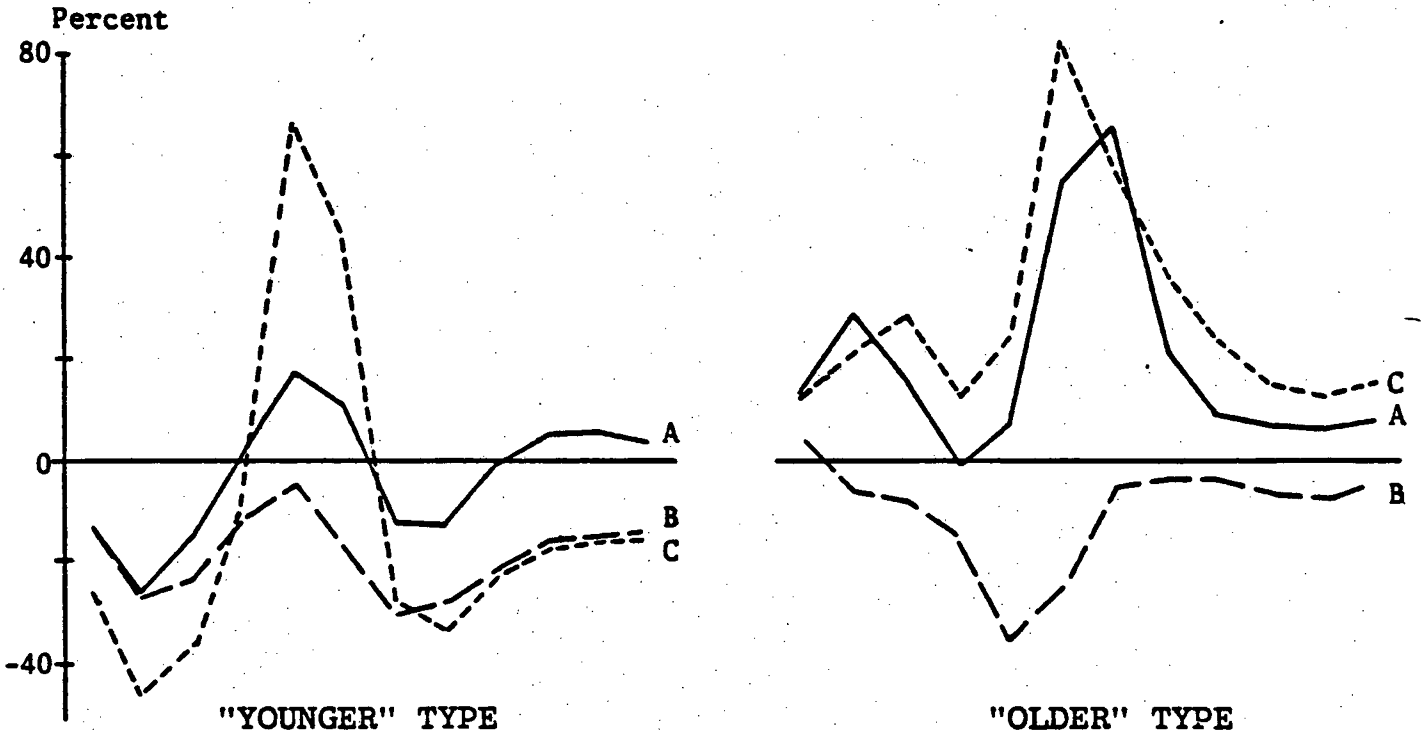
Considerable work remains to be done. The above represents only a first attempt at implementing a computerized, flexible migration model capable of generating large numbers of projections with minimal burden on analytical staff. Several aspects of this model warrant further study; some of the more important 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 model are compared with the county-level race-sex-age population counts from the 1980 Census, can the strengths and weaknesses of this model be fully evaluated. 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.

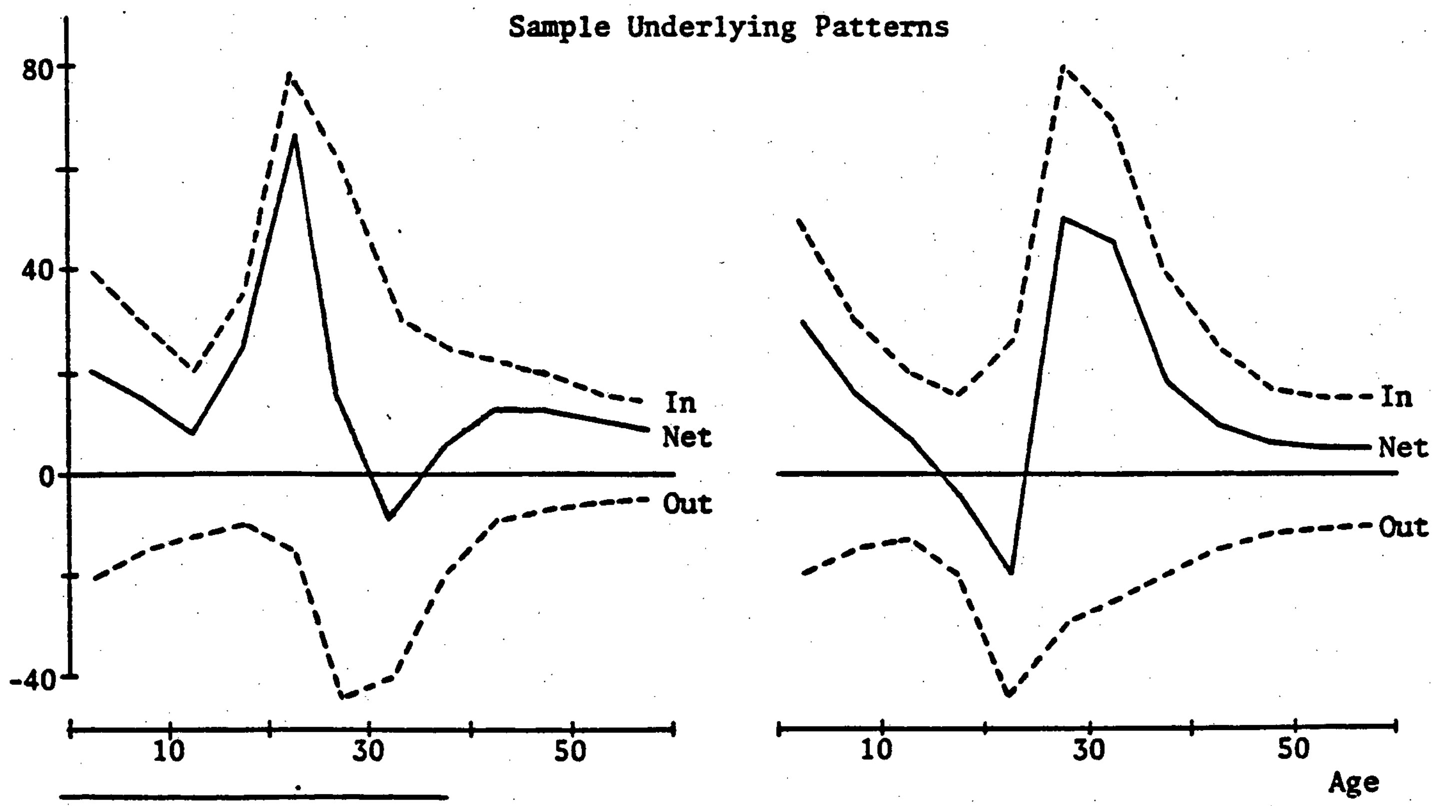
Figure 1.--Illustration of Age-Specific Net Migration Rate Typology Based on Underlying Directional Patterns.



1960-1970 Empirical Examples*

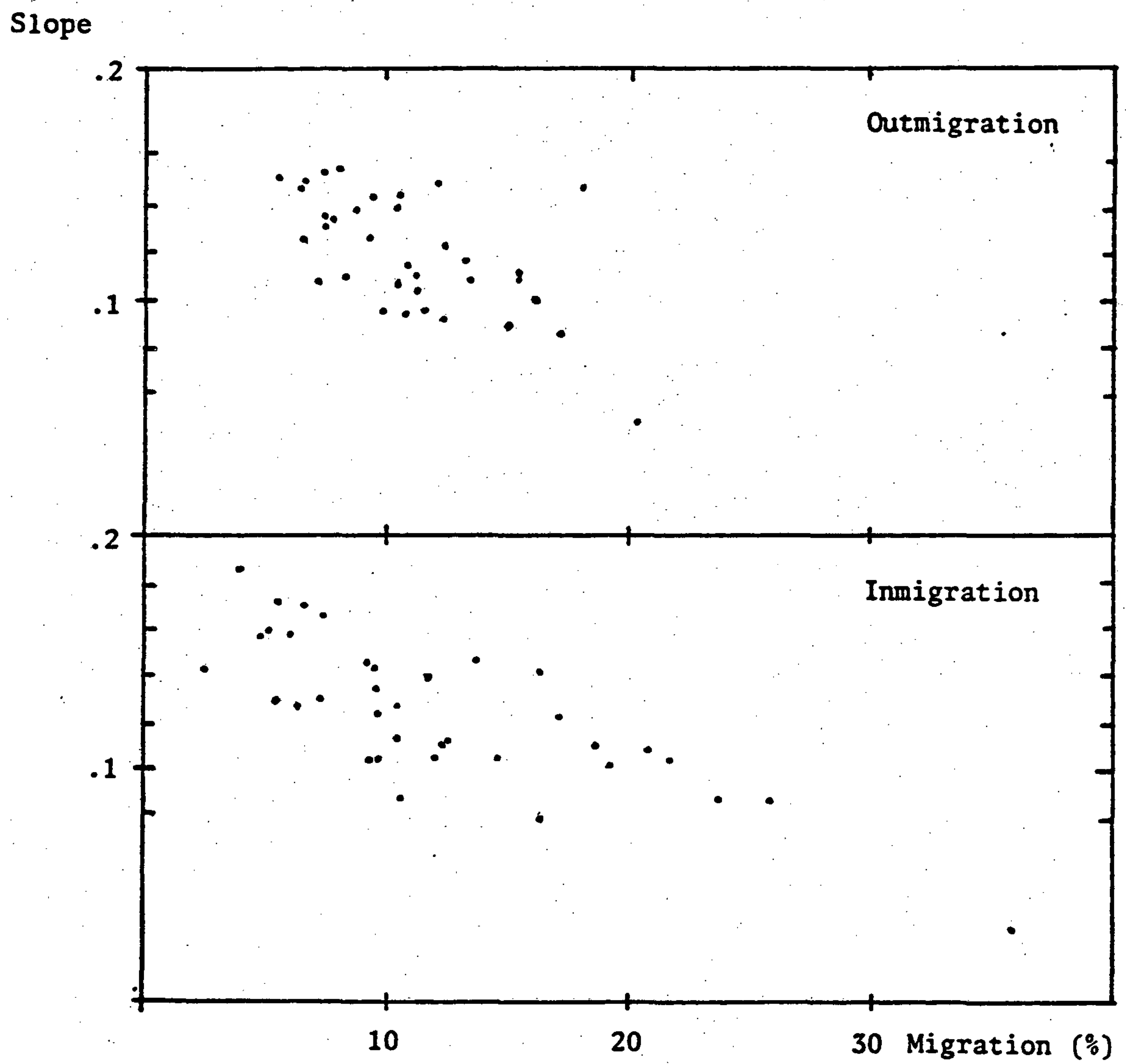
A. Queens Co., N.Y. White Females
B. Kings Co., N.Y. White Females
C. San Francisco Co., Cal. White Females

A. Bucks Co., Pa. White Females
B. Franklin Co., N.Y. Females
C. Burlington Co., N.J. White Females



* Source: Bowles, et. al., 1975.

Figure 2.--Relationship of Migration Slope to Migration Level at Ages 40-44: Selected State Economic Areas, 1965-70; Males.



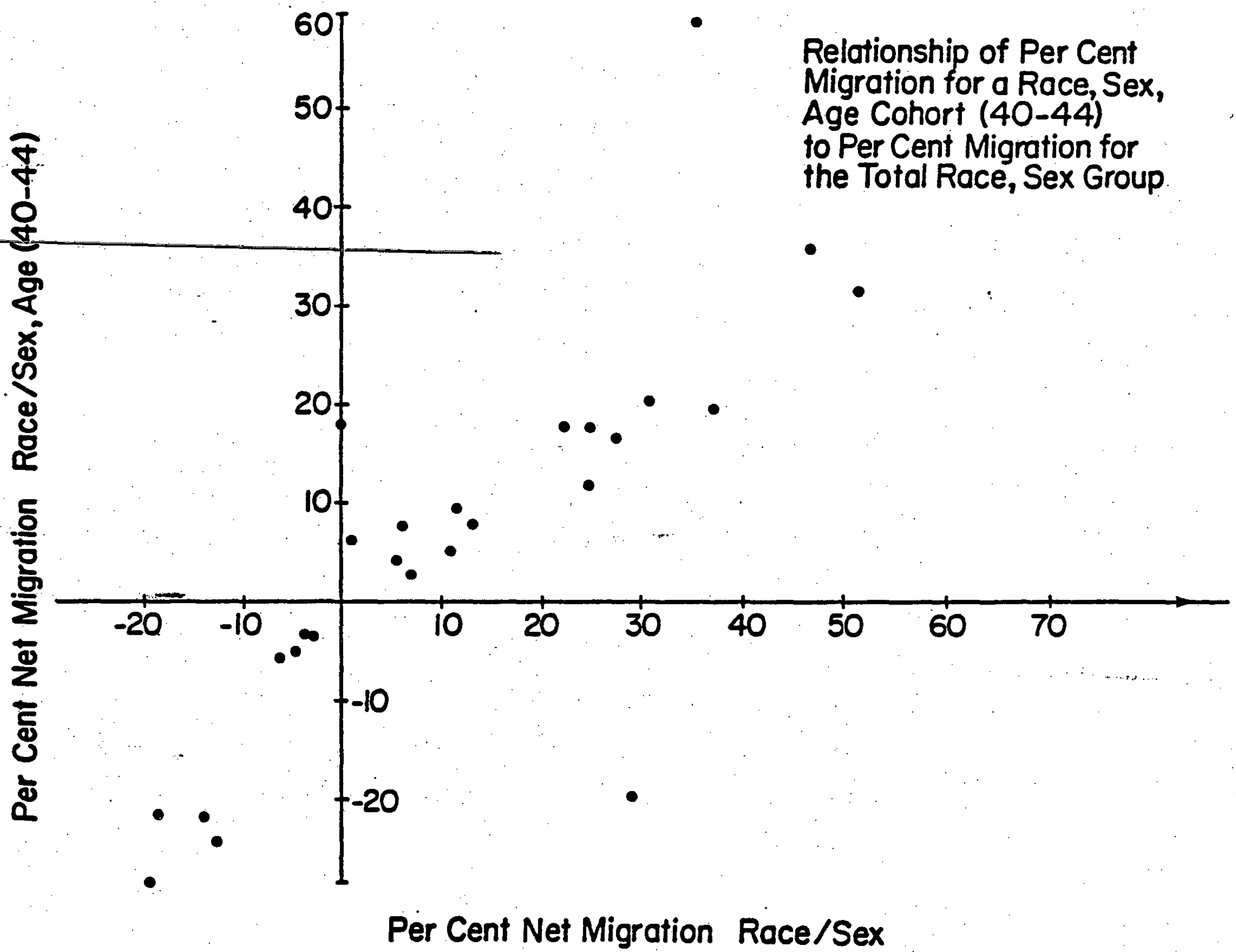
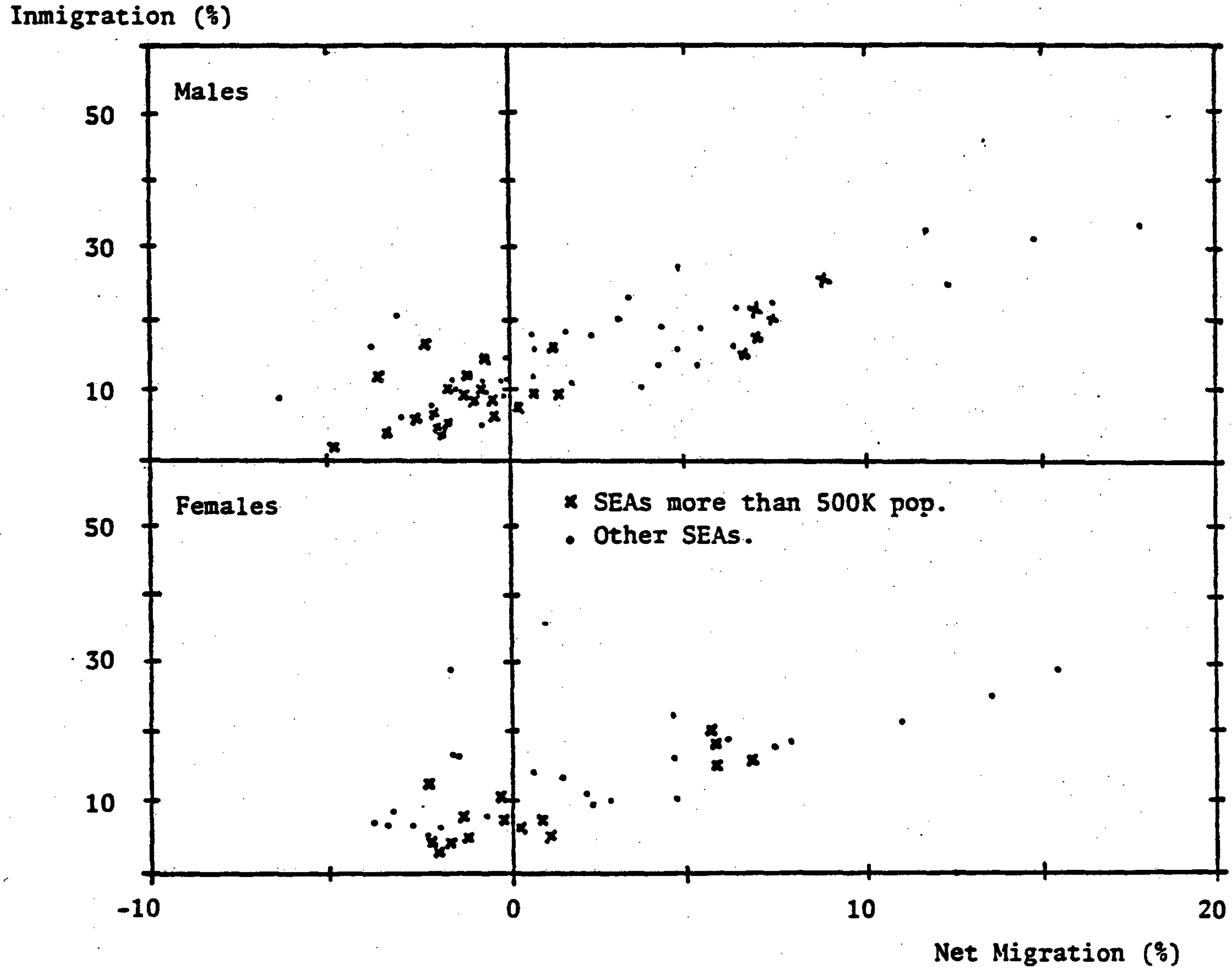

Relationship of Per Cent Migration for a Race, Sex, Age Cohort (40-44) to Per Cent Migration for the Total Race, Sex Group
Per Cent Net Migration Race/Sex, Age (40-44)
Per Cent Net Migration Race/Sex
60
50
40
30
20
10
-10
-20
-20
-10
10
20
30
40
50
60
70

Figure 3

Figure 4 --Relationship of Inmigration to Net Migration, Ages 40-44: Selected State Economic Areas, 1965-70.



1960-1970 Net Migration Rates for Selected Age Cohorts

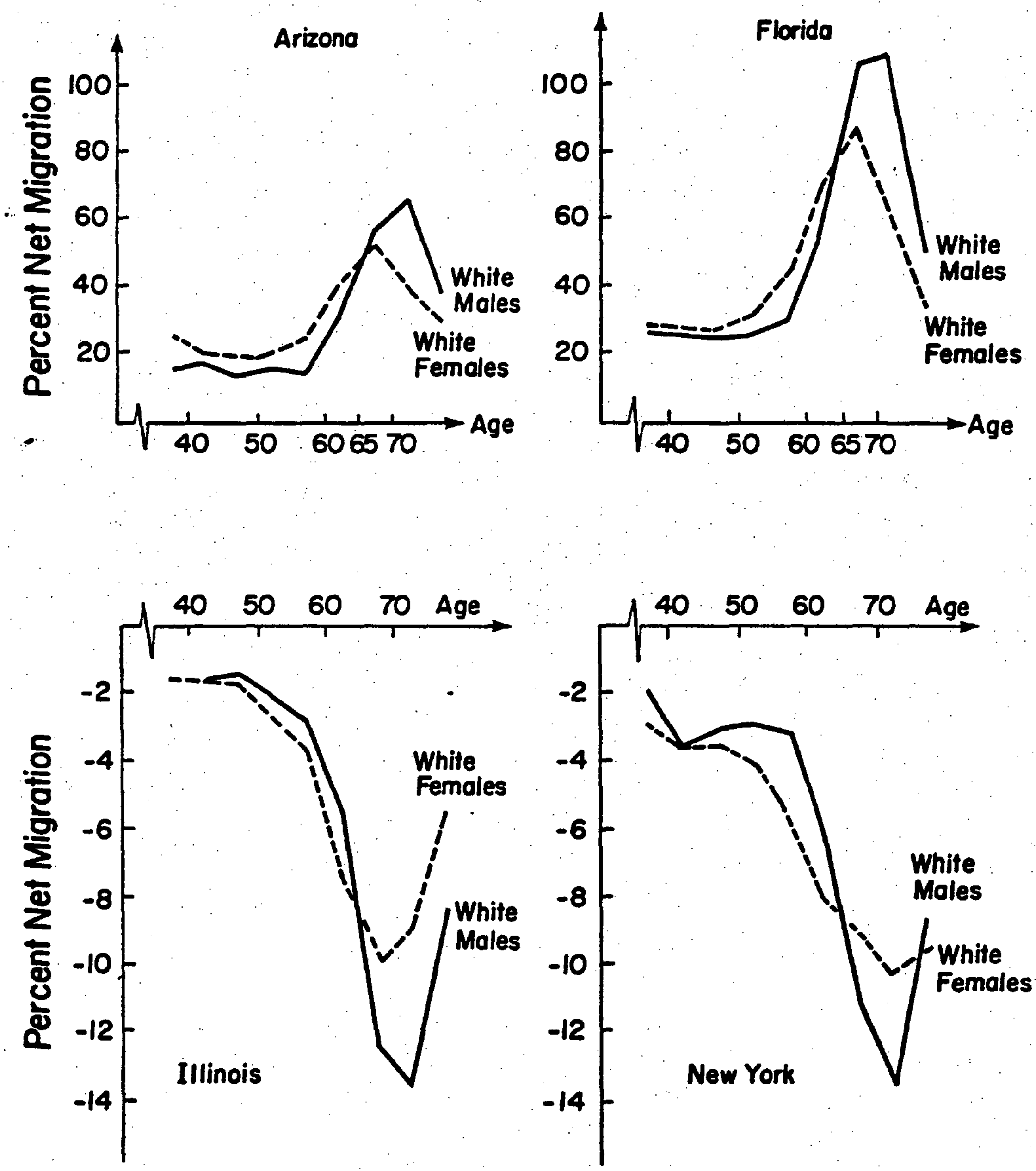
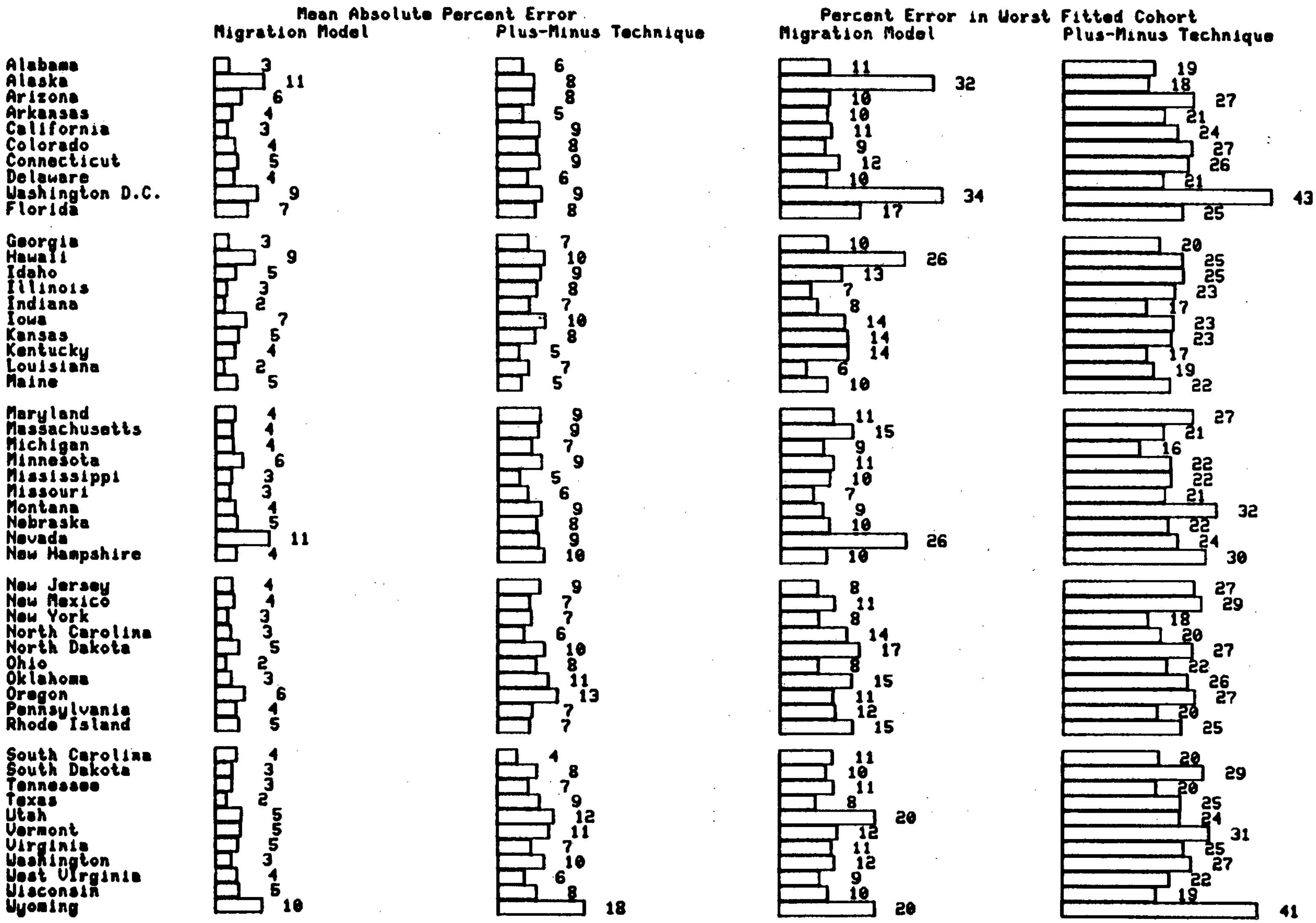


Figure 5

Figure 6
Comparisons of 1980 Population Projections with 1980 Census data



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

Irwin, Richard, 1977, "Guide for Local Area Population Projections", U. S. Bureau of the Census Technical Paper No. 39, Washington D.C., U. S. Government Printing Office

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

----------, 1976, Projecting State and Local Populations. Cambridge, Mass.: Ballinger

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

Rogers, Andrei, 1975, Introduction to Multiregional Mathematical Demography. New York, Wiley-Interscience

Schroeder, Esther, 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, Lawrence Berkeley Laboratory, Berkeley, California

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

VI. References

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

----------, 1979, Current Population Report, Series P-25, Number 796, "Illustrative Projections of State Populations by Age, Race, and Sex: 1975 to 2000", Washington D.C., U. S. Government Printing Office

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

----------, 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", Washington D.C., U.S. Government Printing Office

----------, 1981, Supplementary Reports, 1980 Census Of Population, "Age, Sex, Race and Spanish Origin of the Population by Regions, Divisions, and States: 1980", PC80-S1-1, U.S. Bureau of the Census, Washington D.C., U. S. Government Printing Office

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, $i=1,2,\ldots,16$
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 $i=4,\ldots,7$
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, $i=1,\ldots,16$
IP_i	in migration pattern, $i=1,\ldots,13$
$ØP_i$	out migration pattern, $i=1,\ldots,13$
AIP_i	adjusted in migration pattern, $i=1,\ldots,13$
$AØP_i$	adjusted out migration pattern, $i=1,\ldots,13$
SA_i	student adjustment, $i=5,\ldots,8$
RM_i	retirement migration $i=14,\ldots,16$
RMR_i	retirement migration rate for $i=14,\ldots,16$ 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)