Canadian Studies in Population, Vol. 20(2), 1993, pp. 153-191.

Towards Measuring Uncertainty in Estimates of Intercensal Net Migration

Hallie J. Kintner
General Motors Research Laboratories

David A. Swanson
University of Arkansas at Little Rock

Abstract

We present a system for generating confidence intervals around estimates of intercensal net migration made using the life table survival method. The life table survival method applies a life table to a census count to project survivors at some past or future time points. Net migration is the difference between the projected number of survivors and the enumerated population. Confidence intervals for net migrants in an age-sex group are based on the probability distribution of deaths. They can be adjusted when a life table is unavailable for the population in question. The technique is illustrated with data from Puerto Rico, New Jersey, and Alaska.

Résumé

Nous présentons un système permettant de générer des intervalles de confiance autour des valeurs estimatives de la migration nette intercensitaire à l'aide d'une méthode utilisant la table de survie. Cette méthode applique une table de survie aux résultats d'un recensement afin de projeter le nombre de survivants à un point donné du passé ou de l'avenir. La migration nette est la différence entre le nombre projeté de survivants et la population recensée. Les intervalles de confiance pour la migration nette dans un groupe d'âge-sexe sont fondés sur la distribution des probabilités de décès. Ils peuvent être ajustés quand on ne dispose pas d'une table de vie pour la population en question. La technique est illustrée à l'aide de données venant de Porto Rico, du New Jersey et de l'Alaska.

Key Words: migration, confidence intervals

Introduction

This paper presents a system for developing confidence intervals around estimated net migrants, when the estimates are made using the life table survival method (Shyrock and Siegel, 1976), which has been one of the most widely used methods to estimate net migration (Hamilton and Henderson, 1944; Hamilton, 1961, 1965, 1966; Siegel and Hamilton, 1952; United Nations, 1970). Although confidence intervals have been examined or developed for population projections (Lee, 1974; Lee, 1985; Saboia, 1974; Cohen, 1986; Davis, 1988; Keyfitz, 1981; Voss et al, 1981; Stoto, 1983; Smith, 1988; Alho and Spencer, 1985), for state-level postcensal population estimates (Espenshade and Tayman, 1982) and for substate postcensal population estimates (Swanson, 1989), we are aware of no previous attempts to develop confidence intervals for estimates of intercensal net migration. This method is an extension of a procedure for obtaining confidence intervals for projecting closed group populations (Kintner and Swanson, 1990).

The life table survival method applies a life table to a census count to project survivors at some past or future time points. Net migration is then estimated as the difference between the projected number of survivors and the enumerated population at that time. The confidence intervals developed here take into account random variation in mortality probabilities. The degree of uncertainty in estimates differs according to the number of deaths in the population (size of the area under consideration) and the life table mortality probabilities.

In this paper, we illustrate a system for developing confidence intervals for the number of net migrants specific to an age group obtained from the life table survival method. One advantage of this technique is that the standard error for the group changes systematically with time as this group ages and as the time span between censuses increases. More correctly, this may be termed a standard error that changes over time with a cohort. This is intuitively appealing since we want "uncertainty" to increase as the time span between censuses lengthens. The standard errors for the number of net migrants in each age-gender group may be combined to obtain a standard error for the total number of net migrants.

In using the life table survival method, two situations are commonly found. The first is that a life table for the population in question is constructed from the population in question and the second is that, for whatever reason, a "borrowed" life table is applied to the population in question. While it is possible to develop confidence intervals for net migration, assuming that both the number of deaths on which the life table is based and the life table parameters for the "borrowed" life table hold for the population in question (i.e., it is not "biased"), we also present three methods for adjusting the confidence intervals to reflect the characteristics of the population in question.

When using a "borrowed" life table that is unbiased, relative uncertainty should increase as the size of the population in question decreases (or, more precisely for our system, as the number of deaths decreases) and as the mortality conditions for the population in question differ from those represented by the life table. These adjustments inflate the variance of the estimates to reflect the increased uncertainty. Thus, the first adjustment technique we present is one that controls for differential population size. The second adjustment procedure is for the situation when population size is equivalent but survivorship in the "borrowed" life table is assumed to be a constant multiple of survivorship in the population to which the life table is to be applied. We also suggest a short-cut way to make an adjustment for size in a given application, which may be particularly useful when both the size and survivorship adjustments are needed. The major advantages of these adjustments for borrowed tables are that they broadly cover the most likely situations to be encountered when one must "borrow" a life table. Further, they allow one to measure the impact of alternative assumptions about differential population size and survivorship on confidence intervals when (as is likely the case) the true magnitude of these differentials are only imprecisely known or estimated.

In this paper, we also use the "random error" approach in regard to estimating uncertainty. As a consequence, we also employ the term "confidence interval". We realize that this term has a related but distinctly different meaning in classical inferential statistics than it does in demographic forecasting. From the viewpoint of classical inferential statistics the term "prediction interval" would be closer to our concept of estimating uncertainty than is "confidence interval" (Kmenta, 1971). However, because the term "confidence interval" has gained widespread usage in estimating uncertainty in both demographic forecasting and estimation, we prefer it over "prediction interval".

In presenting this system for placing confidence intervals around estimated net migrants, we find it useful to discuss its limitations by starting with an observation made by Pittenger (1978: 276) about the role of assumptions in assessments of uncertainty in population forecasting:

> Essentially, the confidence intervals are valid only if all of the assumptions in the model application are valid; since the assumptions are judgmental, it follows that the confidence limits are also judgmental.

Pittenger (1978: 272) offers three broad categories of assumptions: (1) strategic; (2) logistical; and (3) tactical. The first, strategic judgment, covers such issues as one's notion of the nature of population change (e.g., deaths occur in a regular manner between two successive census counts) and the methodology that can transform this notion into demographic applications (e.g., life table survivorship values). The second, logistical judgment, involves the selection of specific techniques within a methodological framework (e.g, the decision to use the forward survival life table technique as opposed to the reverse survival life table technique), as well as the selection and use of data. (Should census counts be adjusted for assumed levels of net undercount error?) The third type of judgment, tactical, has to do with the specification of specific values within a technique (e.g., the selection of a given life table, once one has decided to use the forward survival life table technique to estimate net intercensal migration).

The confidence interval procedure that we present for net migration estimates is, like most new procedures, subject to a set of limiting assumptions that we expect will be better understood and, perhaps, relaxed, as more experience is gained with it. At this point, we cite the major assumptions that we have identified in terms of Pittenger's three types.

Most of our "strategic" assumptions are inherent in the life table method itself. However, another important strategic assumption is found in viewing the net number of migrants in each age group as independent variables in order to take advantage of simplified mathematics. In the same vein, we also assume that the shape of the total net migrant distribution is "normal," according to the extended central limit theorem and the assumption that every age group's total follows a normal distribution (Espenshade and Tayman, 1982).

In terms of the "logistical" judgments, the assumptions we make under this category probably play the largest role in the utility of our procedure. For example, we know that in any given use of the life table survival rate method, systematic errors in estimates of net migration exist apart from the stochastic error that can be estimated by our proposed procedure. These errors include differential net census undercount error, death registration error, and "bias" in a given life table. Another includes the fact that the confidence intervals are generated for a particular population for a particular period of time. They may, or may not, apply to the same population at a different point in time or to a different population at the same point in time. In some instances, these sources of error would overwhelm the stochastic error used in our procedure to generate confidence intervals; in others, they would not. In any event, we postpone dealing with systematic error because the level of complexity required is more appropriately dealt with in an article that uses the current paper as a point of departure.

Finally, we have "tactical" assumptions. For example, in the illustrative example applications we provide later in this paper, particular life tables from a set of plausible alternatives were selected.

The main point here is that we present our procedure as one subject to limitations and that as with any demographic application, a user should exercise judgment about its applicability in a given situation. A corollary is that confidence intervals constructed under our procedure are like those constructed for population forecasts: they are valid to the extent that the assumptions within which they are embedded are valid.

Methods of Estimation

Net Migration

Net migration is frequently estimated indirectly because residential (migration) histories are often not available. Hence, net migration is commonly estimated as a residual, the net excess of enumerated population change over that expected under natural increase alone. Consider the problem of estimating the amount of net migration between two time points, with a census taken at each time point. Population change is frequently described by the balancing equation:

$$P(t) = P(0) + (B - D) + (I - E) \tag{1}$$

where the population at time t is viewed as the result of the population at time 0, plus the change from natural increase (the excess of births (B) over deaths (D)) and the change from net migration (the excess of immigrants (I) over emigrants (E)).

The vital statistics method obtains net migration by rearranging terms in the balancing equation to obtain net migration (I-E). It requires complete registration of births and deaths during the intercensal period. Consequently it is inappropriate for use in areas with incomplete vital registration, like less developed countries.

The Survival Rate method (Shryock and Siegel, 1976) is commonly used for indirectly estimating net migration because it does not require complete vital registration. Rather it uses survival rates from a life table or successive censuses. The life table version of the survival ratio method applies a life table to a census count to generate expected survivors at some future or past time point. Net migration is then estimated as the difference between the enumerated population at that past or future time point and the expected number of survivors to that date.

The Forward Survival Method uses a life table to survive the population at the earlier date forward to the later date.

$$NMF(0,t) = P(t) - s \cdot P(0) \tag{2}$$

where NMF(0,t) is the number of net migrants estimated from the Forward Survival method, P(t) is the population at time t, and s . P(0) is the expected number of survivors to time t (from the survival rate s and the population at time 0).

The Reverse Survival Method "reverse-survives" the population at the later time period backwards in time to the earlier time period. Net migration is estimated as the difference between this expected population and the actual population as enumerated in the census. It can be obtained from the previous equation by dividing by the survival rate.

$$NMR(0,t) = \frac{P(t)}{s} - P(0) \tag{3}$$

where NMR(0,t) is the number of net migrants from the Reverse Survival method and the following terms were as previously defined: P(t), s, and P(0).

Although the forward and reverse survival estimation procedures rarely generate identical estimates, they are usually close and, moreover, they are linked by the relationship NMF(0,t) = NMR(0,t)·s.

Another estimate takes the average of the forward and reverse survival estimates.

$$NMA(0,t) = \frac{NMF(0,t) + NMR(0,t)}{2} \tag{4}$$

Again, keep in mind that survival rates for these procedures can come from life tables or from census survival rates. In our system, we utilize survival rates that come from life tables to generate confidence intervals. This approach can therefore be used whenever life tables are available. Note, however, that calculation of standard errors for survivorship ratios assumes that information about the number of deaths by age and gender is available. This may not always be the case in less developed countries where vital registration is inadequate and life tables are estimated indirectly.

Although a point estimate of intercensal net migration is needed for many purposes, it is also important to have a notion of the level of confidence one may have about the precision of the estimate given the set of judgments in which the estimate is embedded. Confidence intervals are constructed by first deciding on the degree of risk one is willing to take of making the error of stating that the point estimate is somewhere in the interval when in fact it is not. For example, with a 95% confidence interval we know that in the long run only 5% of the time would we get intervals by this procedure that would not include the parameter of interest. The confidence interval is obtained by enclosing the point estimate in an interval that is a certain multiple of standard errors corresponding to the degree of risk taken. The confidence interval is bounded by the lower and upper confidence limits.

Confidence Intervals

In a previous paper (Kintner and Swanson, 1990), we developed confidence intervals for projecting survivors from closed group populations. These confidence intervals are based on the statistical properties of the survivorship ratio. In contrast to many other attempts to derive confidence intervals for population projections, the confidence intervals developed for surviving closed populations are based on the probability distribution of deaths, given that the mortality structure used to project survivors will remain unchanged (Smith, 1988; Stoto, 1983).

The stochastic aspect of life table parameters is frequently ignored, although the mortality rates underlying these parameters are subject to random variation, as Chiang (1984, p.78) has pointed out.

> Statistically speaking, human life is a random experiment and its outcome, survival or death, is subject to chance. If two people were subjected to the same risk of dying (force of mortality) during a calendar year, one might die during the year and the other survive. If a person were allowed to relive the year he survived the first time, he might not survive the second time. Similarly, if a population were allowed to live the same year over and over again, the total number of deaths occurring the second time would assume a

different value and so, of course, would the corresponding death rate. It is in this sense that a death rate is subject to random variation even though it is based on the total number of deaths and the entire population.

Chiang has derived formulas for the sample variances of life table functions. In our previous paper we extended Chiang's formula for the variance of the probability of surviving from age i to age i+1 to the survivorship ratio (the probability of surviving from age i+0.5 to age i+1.5) (Kintner and Swanson, 1990). We are able to generate confidence intervals for projections of populations closed to renewal and to decrements other than mortality from the variance of the survivorship ratio. Confidence intervals are obtained by enclosing the point estimate in an interval that is a certain multiple of the standard error, with the multiplier related to the degree of risk taken.

As was mentioned earlier, all the life table survivorship methods estimate intercensal net migration as the difference between the actual population at a date and the projected number of survivors to that date. For instance, the Forward Survival Method projects the population at the earliest date forward to the latest date by applying the survivorship ratio. If we assume that the population counts at both census dates are constants, then the variance of NMF is merely a function of the variance of the survivorship ratio.

$$\begin{aligned} Var(NMF(0,t) &= Var(P(t) - s \cdot P(0)) \\ &= Var(P(t)) + Var\,(s \cdot P(0)) \\ &= P^2\,(0) \cdot Var(s) \end{aligned}$$

Similarly, the Reverse Survival Method projects a population back in time (called backcasting) by dividing the population count at the latest date by the survivorship ratio. We obtain the variance for the inverse of the survivorship ratio by approximating the inverse by the sum of a convergent power series. Finally, we obtain the variance for the Average of the Forward and Reverse Survival Methods.

The procedures described here provide confidence intervals for the number of net migrants in an age group. Confidence intervals for total net migrants are based on the variance of total migrants, which can be obtained by summing the variances of the age groups in most applications. This sum is appropriate in the examples presented here because the net migrants in the age groups are independent variables. Independence holds when the survivorship ratios do not refer to overlapping age groups. Under other circumstances the variance for total net migrants is the sum of the variances of the age groups and the covariance between adjacent age groups; for further details see Kintner and Swanson, 1990.

Appendix A presents the mathematical basis of the confidence intervals for estimating net intercensal net migration based on life table survival rate methods.

An important limiting assumption mentioned earlier as one of our "strategic" judgments, is that the life table used to describe survival remains in effect over the time between censuses (i.e., the pattern of age-specific mortality reflected in the calculated life table is invariant over the time frame of this study).

In reality, this assumption is often violated. Historically, survivorship has improved over long time periods. However, it is also true that this assumption is widely used in the forecasting or projecting of populations as well as in obtaining estimates of net migration: a life table for a specific time period is usually assumed to hold over the life span of a generation.

Applications

Puerto Rico, 1950-1960: No Adjustments

We first illustrate this procedure by applying it to the problem, presented in Shyrock and Siegel (1976), of estimating the number of male net migrants by age for Puerto Rico, 1950-1960. Population counts and 10 year survivorship ratios based on the 1954-1956 Puerto Rican life table come from Shyrock and Siegel Table 20-10 (p. 595). We obtained deaths by age and sex for Puerto Rico 1954-1956 from the United Nations Demographic Yearbooks (United Nations, 1957 and 1962).

Table 1 presents the results. The Forward Survival method estimates that outmigration exceeded inmigration by a total of (-)258,542 males between 1950 and 1960. We find that the 95% confidence interval for this estimate is between -259,232 and -257,852. The absolute width of the confidence interval is 0.27% of the estimated net number of migrants.

The Reverse Survival Method estimates that outmigration exceeded inmigration by (-)265,856 males. The 95% confidence interval for this estimate is between -266,806 and -264,906. The absolute width of the confidence interval is 0.36% of the estimated net number of migrants.

TABLE 1. 95% CONFIDENCE INTERVALS FOR ESTIMATES OF MALE NET MIGRANTS FOR PUERTO RICO, 1950-1960

Age 1950	*in 1960*	*Population 1950*	*in 1960*	*10-year Surv. Ratio*	*Forward Estimate*	*Method 95%CI*	*Backward Estimate*	*Method 95%CI*	*Average Estimate*	*Method 95%CI*
Births,1955-60*	0-4	196,140	179,619	.93747	-4,256	±251	-4,540	±261	-4,398	±256
Births,1950-55+	5-9	206,277	165,930	.92024	-23,894	±272	-25,965	±249	-24,930	±263
0-4	10-14	185,014	162,244	.97658	-18,437	±107	-18,879	±75	-18,658	±98
5-9	15-19	161,446	122,602	.99004	-37,236	±94	-37,611	±73	-37,424	±83
10-14	20-24	138,696	79,792	.98717	-57,125	±130	-57867	±77	-57,496	±103
15-19	25-29	108,984	61,971	.97984	-44,816	±145	-45,738	±86	-45,738	±115
20-24	30-34	91,269	58,723	.97361	-30,137	±152	-30,954	±103	-30,954	±128
25-29	35-39	76,528	61,592	.96898	-12,562	±137	-12,964	±118	-12,964	±127
30-34	40-44	66,769	53,087	.96148	-11,110	±125	-11,555	±107	-11,555	±116
35-39	45-49	67,324	53,781	.95196	-10,309	±156	-10,829	±137	-10,829	±147
40-44	50-54	47,745	39,832	.94135	-5,113	±128	-5,431	±120	-5,431	±124
45-49	55-59	39,893	34,404	.92389	-2,453	±128	-2,655	±129	-2,655	±129
50-54	60-64	36,548	29,095	.88796	-3,358	±151	-3,782	±152	-3,782	±152
55-59	65-69	24,692	24,525	.81677	+4,357	±144	+5,335	±215	+5,335	±179
60-64	70-74	25,636	16,370	.71369	-1,926	±206	-2,699	±259	-2,699	±232
65-69	75-79	16,270	10,275	.61203	+317	±171	+518	±289	+518	±230
70-74	80-84	10,679	4,636	.52046	-922	±138	-1,771	±220	-1,771	±178
75+	85+	13,453	4,286	.28604	+438	±175	+1,531	±680	+1,531	±427
All ages	All	1,110,946	1,162,764		-258,542	±690	-265,856	±950	-262,201	±765

Source: Shyrock and Siegel, p. 595.

*Equals three-fourths of the births in 1955 plus the births in 1956, 1957, 1958, 1959, and one-fourth of the births in 1960.

+Equals three-fourths of the births in 1950 plus the births in 1951, 1952, 1953, 1954, and one-fourth of the births in 1955.

The Average method finds that there were (-)262,201 more outmigrants than inmigrants. The 95% confidence interval for this estimate is between -262,966 and -261,436. The confidence interval width is 0.29% of the estimated net number of migrants.

One advantage of this technique is that it generates confidence intervals for the number of net migrants in an age-gender category. Examination of the absolute and proportionate size of the confidence intervals for the age groups indicates that the absolute width of an interval is largest where there are the highest numbers of deaths. The interval width exceeds 5% of the estimate in the following age groups regardless of estimation method: less than 5, 70-74, 75-79, 80-84, and 85+ (ages are those in 1960).

A few notes about how the confidence intervals for Puerto Rico were estimated, in case any readers wish to do the calculations themselves. First, an abridged life table is used although the derivations in the appendix are in terms of single year age groups. So, the formulas in the appendix were changed to increment by five years rather than by a single year. Second, Shyrock and Siegel (1976) used a life table presented in the United Nations Demographic Yearbook (1961). As they note, they calculated survivorship ratios from L_x values obtained by differencing T_x rather than from l_x. The formula for the variance of the survivorship ratio requires estimation of $q_{i+2.5}$, which is estimated by $(0.5*(d_i+d_{i+5})/L_i$ in the appendix. Here, however, we estimated $q_{i+2.5}$ by the converse of the five year survivorship ratio $(1-(L_{i+5}/L_i)$ for ages less than 85 in 1960, and by $(1-(T_{85}/T_{75})$ for the open-ended age group. Third, following Shyrock and Siegel (1976), survivorship of the age cohorts born during the intercensal period (and thus exposed to the risk of dying for less than ten years) is treated specially. The survival ratio for the cohort ages 0-4 in 1960 (births, 1955-60) is $L_{0\text{-}4}/5l_0$ and that for the cohort ages 5 to 9 in 1960 is $L_{5\text{-}9}/5l_0$.

New Jersey, 1960-1970: Adjusting for Size in a "Borrowed Life Table"

The second application estimates net migration for the state of New Jersey for 1960-70 using the reverse survival method. This application demonstrates how confidence intervals for net migration can be adjusted when using a "borrowed" life table that corresponds to the life table in question but is taken from data representing a larger population. Since (hypothetically) no life table is available for New Jersey, survivors backcast from 1970 have been estimated from a life table for the Mid-Atlantic region (New Jersey, New York and Pennsylvania). This specific example represents a typical situation found in practice: net numbers of migrants are estimated for a given area using a life table constructed for a larger geographic area that includes the area in question.

In this specific example, an abridged 1970 life table for the male population of the Mid-Atlantic region (MAD) was constructed using Greville's method (Shyrock et al, 1976, p. 255). Age-sex specific death rates of the whole region for 1970 were calculated and then used to construct the life table for males. The calculation of death rates by age and sex for MAD is based on available deaths by age and sex for the three states which were drawn from U.S. Vital Statistics (United States, 1960-1973) and from counts of persons by age and sex in 1970 (United States, 1973). Table 2 presents the life table for males in the Mid-Atlantic Division, 1970.

TABLE 2. LIFE TABLE FOR MALES, MID-ATLANTIC DIVISION, 1970

Age Group	q_x	l_x	d_x	L_x	T_x	e_x
0	.02332	100,000	2,332	98,438	6,710,930	67.1
1-4	.00328	97,668	320	390,244	6,612,492	67.7
5-9	.00035	97,348	34	485,714	6,222,248	63.9
10-14	.00404	97,314	393	485,185	5,736,534	58.9
15-19	.00713	96,921	691	483,217	5,251,349	54.2
20-24	.01050	96,230	1,010	478,673	4,768,132	49.5
25-29	.00867	95,220	826	474,713	4,289,459	45.1
30-34	.01218	94,394	1,150	469,388	3,814,746	40.4
35-39	.01593	93,244	1,485	462,619	3,345,358	35.9
40-44	.02305	91,759	2,115	453,863	2,882,739	31.4
45-49	.03580	89,644	3,209	440,797	2,428,876	27.1
50-54	.05676	86,435	4,906	420,755	1,988,079	23.0
55-59	.09173	81,529	7,479	390,141	1,567,324	19.2
60-64	.13075	74,050	9,682	347,400	1,177,183	15.9
65-69	.19749	64,368	12,712	291,426	829,783	12.9
70-74	.26519	51,656	13,699	225,016	538,357	10.4
75-79	.35934	37,937	13,639	155,892	313,341	8.3
80+	1.0000	24,318	24,318	157,449	157,449	6.5

Table 3 presents estimates of net migration in New Jersey using the reverse survival method based on the 1970 life table for the Mid-Atlantic Division. Table 3 indicates that there were 227,300 more in-migrants to than out-migrants from New Jersey during the period 1960-1970. This table also shows unadjusted and adjusted confidence intervals for the estimates of net migration. The unadjusted confidence interval is ±307 persons.

We derive (in appendix B) a method to adjust the variance of the estimates of net migration (and hence the confidence intervals) when no life table is available for the population in question so a life table is "borrowed" from a population different in size. It is assumed that the life table from the other population applies to the population in question. Since the population in question is often smaller than the population from which the life table was obtained, the adjusted confidence intervals reflect the smaller number of deaths of the population in question.

The three columns on the right hand side of Table 3 show the results. The third column from the right displays the actual number of deaths in the Mid-Atlantic Division to men in each age group. This column provides the deaths used to calculate the life table in Table 2. The second column from the right shows the number of deaths in New Jersey to men in each age group. These deaths were used in place of the deaths in the Mid-Atlantic Division to calculate an adjusted variance for the number of net migrants in each age group. The right hand column shows the 95% confidence intervals adjusted for the difference in the number of deaths between New Jersey and the Mid-Atlantic Division. When this adjustment is made, the confidence intervals become much wider than the unadjusted intervals. For instance, the confidence interval for the total number of net migrants grows from ±307 to ±725. The absolute width of the confidence interval increases from 0.14% to 0.32% of estimated total net migrants.

Fairbanks-North Star Borough, Alaska, 1970-1980: Adjusting for Both Size and Survivorship Differences in a "Borrowed" Life Table

This application involves the estimation of net migration for Fairbanks-North Star Borough, Alaska between 1970 and 1980 using the reverse survival method. We chose this county because, although it is one of the largest in Alaska, it is sufficiently small to be typical of estimates made by state and local demographers. The size of the borough was 45,864 in 1970 and 53,983 in 1980, according to the respective censuses. The application demonstrates how confidence intervals for net migration can be adjusted for both size and differential survivorship when using a "borrowed" life table that is systematically different than the life table for the population in question. Since no life table is available for Fairbanks-North Star Borough during the period in question, 1970 survivors backcast from 1980 have been estimated using a 1970 Alaska state life table (Alaska, 1983).

Here we are simulating the following situation: in order to estimate net numbers of migrants we must use a life table that represents a systematically

TABLE 3. UNADJUSTED AND ADJUSTED 95% CONFIDENCE INTERVALS FOR REVERSE SURVIVAL ESTIMATES OF MALE NET MIGRANTS FOR NEW JERSEY, 1960-1970

Age 1960	*in 1970*	*Population 1960*	*Population 1970*	*10-yr. Surv.Ratio*	*Net Migrants*	*Unadj. 95%CI*	*MAD Deaths*	*NJ Deaths*	*Adj. 95%CI*
Births,1955-60	0-4	340,339	300,530	.97736	-32,847	±26	8,060	1,548	±59
Births,1950-55	5-9	302,671	354,242	.97143	61,989	±43	803	148	±97
0-4	10-14	326,239	361,478	.99284	37,846	±32	805	154	±72
5-9	15-19	296,623	310,254	.99486	15,234	±20	2,185	375	±47
10-14	20-24	266,974	241,859	.98658	-21,825	±22	2,838	469	±53
15-19	25-29	199,252	221,527	.98240	26,244	±23	2,240	378	±56
20-24	30-34	156,024	194,658	.98060	42,485	±23	2,189	409	±55
25-29	35-39	174,957	200,311	.97452	30,591	±29	3,128	561	±67
30-34	40-44	210,829	225,932	.96692	22,833	±38	5,283	973	±88
35-39	45-49	228,632	229,822	.95283	12,567	±45	8,351	1,600	±104
40-44	50-54	217,076	210,840	.92705	10,355	±53	12,286	2,325	±102
45-49	55-59	200,775	182,908	.88508	5,882	±65	17,762	3,334	±149
50-54	60-64	173,641	147,723	.82566	5,274	±75	22,979	4,010	±175
55-59	65-69	149,927	108,674	.74698	-4,443	±63	26,116	4,616	±200
60+	70+	371,042	176,615	.45733	15,145	±207	94,155	16,296	±608
All ages	All ages	3,615,001	3,467,373		227,330	±307	115,025	37,196	±725

different survivorship function than does the "true" but unavailable life table. An example of such a situation would be the application of a level 18 "West Model Life Table" to a population that is at level 14 of the same family.

Table 4 presents reverse survival estimates of 1970-1980 net migration for Fairbanks-North Star Borough using the 1970 Alaska state life table. It shows that there were 1,288 more in-migrants to than out-migrants from Fairbanks-North Star during the period 1970-1980. We find that the unadjusted 95% confidence interval for this estimate is between 1,267 and 1,309. The absolute width is 1.63% of the estimated net number of migrants, much higher than the respective percentage in the other examples because of the relatively small size of the borough.

These unadjusted confidence intervals assumed that it was appropriate to borrow the state life table without any modifications. Next, we present confidence intervals adjusted both for the smaller population size and for an assumption that survivorship in Fairbanks-North Star Borough is better than in the entire state. In Appendix B, we detail the derivation of the method used to adjust the variance of the estimated net number of migrants when a "borrowed" life table with differential survivorship is applied to the population in question. It is useful to note that in 1970, life expectancy at birth in Alaska was 70.6 years; by 1980, it increased to 72.7 (Alaska, 1983). In this illustrative example, we assume that the pattern of survivorship in Fairbanks-North Star is similar to the pattern for the State of Alaska and that mortality differences at any given age between the two populations can be approximated by a constant multiple.

In our example, we assume that age-specific mortality rates in Fairbanks-North Star Borough are 80% of those in the state of Alaska. To obtain confidence intervals adjusted for both size and survival differences, we first estimate the number of deaths in an particular age group in the borough from the number for the state (shown in the third column from the right in Table 4). We assumed that deaths in the borough occur at 80% of the state age-specific mortality rate (results shown in second column from the right in Table 4). Then we adjust the confidence intervals using the same formulas as for the preceding example.

An alternative procedure, which adjusts only for survivorship differences, is to first find the mean of the q_x values from the 1970 state life table (excluding q_x for the final, open-ended age group) and then "estimate" the mean of the Fairbanks-North Star q_x values as about 80% of the state mean value.

TABLE 4. UNADJUSTED AND ADJUSTED 95% CONFIDENCE INTERVALS FOR REVERSE SURVIVAL ESTIMATES OF NET MIGRANTS FOR FAIRBANKS-NORTH STAR BOROUGH, ALASKA, 1970-1980

Age 1970	*Age 1980*	*Borough Population 1970*	*Population 1980*	*State 10-year Surv.Ratio*	*Net Migrants*	*Unadj. 95% CI*	*State AK Deaths*	*Borough Est. Deaths*	*Adj. 95% CI*
Births,1975-79	0-4	6,020	5,804	.97718	-80	±3	190	35	±7
Births,1970-74	5-9	6,370	4,945	.97289	-1,287	±4	23	2	±9
0-4	10-14	4,817	4,378	.99104	-399	±3	13	2	±7
5-9	15-19	5,115	4,964	.99049	-103	±2	53	8	±7
10-14	20-24	4,501	7,507	.98489	3,121	±5	77	13	±13
15-19	25-29	4,111	7,552	.97901	3,603	±6	59	14	±13
20-24	30-34	8,819	6,431	.97841	-2,246	±5	46	10	±11
25-29	35-39	4,680	4,475	.97695	-99	±4	63	11	±9
30-34	40-44	3,502	3,084	.96952	-321	±4	77	11	±9
35-39	45-49	3,083	2,363	.95379	-606	±4	108	14	±10
40-44	50-54	2,263	1,984	.93483	-141	±4	99	14	±9
45-49	55-59	1,909	1,582	.93884	-224	±4	109	16	±12
50-54	60-64	1,490	1,053	.86852	-278	±5	118	18	±12
55-59	65-69	1,045	953	.77329	187	±8	109	26	±18
60+	70+	1,165	725	.54662	161	±13	165	38	±24
All ages	All ages	58,890	57,800		1,288	±21	1,309	232	±47

Sources: Alaska, 1983.

The two columns to the right in Table 4 show the unadjusted confidence intervals and the confidence intervals adjusted for the survivorship difference. When the adjustment is made, the confidence intervals become much wider than the unadjusted intervals. The confidence interval for total net migrants grows from ±21 to ±47. The absolute width of the adjusted confidence interval is 3.8%.

In Table 4 the adjusted confidence interval is 224% of the unadjusted confidence interval. We can distinguish the impact on the interval width of the survivorship difference separately from the size difference. Appendix B shows that if the confidence interval had been adjusted for differential survivorship only, the adjusted interval would be 123.2% of the unadjusted interval. We can conclude that the adjustment for differential survivorship has slightly more impact than does the adjustment for differential size.

Conclusion

This paper extends the methodology for estimating the number of net migrants by providing a system for generating confidence intervals around such estimates. There is a sound theoretical foundation for deriving confidence intervals around estimates of net intercensal migration obtained from the three survival rate methods. These methods all estimate net migration as the difference between the enumerated population at a date and the projected number of survivors to that date, where the projection is based on a life table applied to a population closed to decrements other than death and to increments other than (known) intercensal births.

The confidence intervals are embedded in a set of strategic, logistical, and tactical judgments. In our illustrative applications, for example, we make the logistical assumption that the population counts in an age-gender group at the beginning or end of the period are known and the strategic assumption that a single life table remains in effect during the entire projection period. This assumption is violated in many applications, but as an initial assumption in the development of confidence intervals it offers the rather strong advantages described in the introduction. Here again, recall that this assumption is one that we plan to deal with in a subsequent paper that uses this one as a point of departure.

We first illustrated our system using the "non-adjustment" technique which assumes that the size and survivorship of the population used to construct the life table are equivalent to the population to which the life table is applied to estimate net numbers of migrants. The application concerns estimating net migrants for Puerto Rico 1950-1960. The confidence intervals for the total number of net migrants are relatively narrow (less than one percent). However, the interval width does differ by age group such that it is widest for the age groups with the largest number of deaths. In the case of Puerto Rico, the youngest and oldest age groups had the proportionately largest confidence intervals.

We have also provided proofs for two types of adjustment to the confidence interval to allow for the situation where a life table for the population in question is not available and, instead, survivors are projected using a borrowed life table. These straightforward extensions greatly expand the range of situations in which confidence intervals can be provided for estimates of net migration.

We then illustrated the adjustment procedure for differential population size by estimating unadjusted and adjusted confidence intervals for the number of net migrants to New Jersey during 1960-1970. The life table for the Mid-Atlantic Division was "borrowed" to estimate net migrants for New Jersey using the reverse survival method. The confidence intervals were adjusted to take into account the smaller number of deaths in New Jersey than in the Mid-Atlantic Division. The adjusted confidence intervals are much wider (roughly double) the unadjusted confidence intervals.

In the illustration of the adjustment for both size and survivorship differences, we used a 1970 Alaska state life table to estimate net migration in the Fairbanks-North Star Borough for the period 1970-1980. Both the unadjusted and adjusted confidence intervals are much wider for the small borough than they are for the other examples based on larger populations. The adjusted intervals are much wider (224%) than the unadjusted confidence intervals.

In Appendix B, we suggest a short-cut way that both adjustments can be made in a given application. We do not illustrate this technique.

Confidence intervals for net migration estimates are useful for detailed planning. For example, when such confidence intervals are wide, it suggests that one needs the ability to shift resources quickly to respond to migration flows. Confidence intervals would also be useful for profiles of net migrants, which could be incorporated into population projections (Pittenger, 1976).

We conclude by noting, again, that our approach is subject to limitations. Further work will result in a greater understanding of these limitations and their effects. It will also likely lead to extensions of this work, including the ability to deal with random mortality variation over time as well as by age. Even with its limitations, our procedure allows one to make formal statements about error, which in the area of population estimation is not often found. The importance of this step is, perhaps, best summarized by Rives (1982: 85), who, in arguing for his proposed "survey-based" method, states:

> "Most population estimation techniques, particularly demographic techniques, permit only statements of error that tend to be judgmental in nature. Such statements can be useful, but they do not always have a strong empirical basis."

Appendix A - Derivation of Confidence Intervals for the Number of Net Migrants by Age

This section derives confidence intervals for the number of net migrants age i (at time 0) who migrate between time 0 and time t. It considers three estimation methods: forward survival, reverse survival and the average of the forward and reverse survival methods.

Both the forward and reverse survival method consider the projected population a function of the population at one time-point (the jump-off population) and a Leslie-type matrix representing the probability of surviving between that time and the date of the projected population for each age group separately by gender. By definition, the calculation of confidence intervals requires estimating variance. We assume that the number of persons in an age-sex group at each census date is known (observed without any sampling error). Consequently, this section focuses on deriving the variance of the survivorship matrix. In this section, we assume that the survival rate is the survivorship ratio from a life table L_{i+1}/L_i.

Throughout the appendix we assume, for ease of notation, that the population in an age-gender group subject to the survival rate was enumerated in both censuses. In most situations, however, the youngest individuals in the latest census were too young to have been enumerated in the earliest census (because they had not yet been born). The number of births during the time period corresponding to the youngest age group(s) would then replace the population counts. Despite this substitution, we retain the general notation to facilitate presentation.

The Forward Survival Method

Net migration is estimated as the difference between the enumerated population at a later time point and the projected number of persons surviving from the earlier census date to the later date.

$$NMF_i(0,t) = P_{i+t}(t) - S_i \cdot P_i(0) \qquad (5)$$

where $NMF_i(0,t)$ is the number of net migrants between age i at time 0 who migrated between years 0 and t, $P_{i+t}(t)$ is the population age (i+t) at time t, and $S_i \cdot P_i(0)$ is the projected number of persons age i at time 0 surviving to age (i+t) at time t (projected from the survivorship ratio S_i and the population age i at time 0, $P_i(0)$). This section derives confidence intervals for the number of net migrants in a single age-gender group for censuses that are one year apart and then extends these results to the more general situation where the censuses are t years apart.

First, consider estimating net migration from censuses that are one year apart. For a single age-gender category i, let

$$NMF_i(0,1) = P_{i+1}(1) - S_i \cdot P_i(0) \tag{6}$$

where: $NMF_i(0,1)$ is the number of net migrants age i at time 0 who migrate between time 0 and time 1, $P_{i+1}(1)$ is the number of persons age i+1 at time 1, and $P_i(0)$ is the number of persons age i at time 0, and S_i is the survivorship ratio from the life table (L_{i+1}/L_i).

We want confidence intervals for $NMF_i(0)$ in the following form:

$$\left(N\hat{M}F_i(0,1) - 1.96\sigma_{N\hat{M}F_i(0,1)}\right) \langle NMF(0,1) \langle \left(N\hat{M}F_i(0,1) + 1.96\sigma_{N\hat{M}F_i(0,1)}\right) \tag{7}$$

where: $N\hat{M}F_i(0,1)$ is the estimated number of net migrants age i at time 0 who migrate between time 0 and time 1, $NMF_i(0,1)$ is the true number of net migrants age i at time 0 who migrate between time 0 and time 1, $\sigma_{N\hat{M}F_i(0,1)}$ is the standard deviation of $N\hat{M}F_i(0,1)$, and 1.96 is the value of the Z test statistic at the 0.05 level. We assume the normal approximation to the binomial distribution to use the Z test. We make this assumption because statistical tables of the critcal values of the Z test are widely available, and to maintain consistency with Chiang. It is also possible to approximate the binomial with the Poisson distribution (Ross, 1989) using various statistical packages. The assumption of the normal approximation is used only for statistical inference.

Next we derive the variance of $NMF_i(0,1)$ and its estimator,

$$\sigma^2_{N\hat{M}F_i(0,1)} = P_i(0)^2 \cdot \frac{\hat{q}^2_{i+.5}\left(1-\hat{q}_{i+.5}\right)}{D_{i+.5}} \tag{8}$$

where $\hat{q}_{i+.5}$ is the probability of dying between exact ages i+0.5 and i+1.5, $D_{i+.5}$ is the number of deaths that actually occurred between ages i+.5 and i+1.5 (as distinct from $d_{i+.5}$, which is the number of deaths in the life table between those ages), and $P_i(0)$ is the jump-off population age i. The proof follows.

From equation (6):

$$\begin{aligned} Var(NMF_i(0,1)) &= Var(P_{i+1}(1) - S_i\, P_i(0)) \\ &= Var(P_{i+1}(1)) + P_i(0)^2\, Var(S_i) \\ &= P_i(0)2\, Var(S_i) \end{aligned}$$

We assume that the population counts at both census dates are constants.

Next we derive $Var(S_i)$ in terms of life table parameters. By definition, for all ages above age one,

$$Var(s_i) = Var\left(\frac{L_{i+1}}{L_i}\right)$$

$$= Var\left(\frac{.5(l_{i+1} + l_{i+2})}{.5(l_i + l_{i+1})}\right)$$

$$= Var\left(\frac{l_{i+1.5}}{l_{i+0.5}}\right)$$

$$= \sigma^2_{p_{i+0.5}}$$

where l_i is the number of survivors to age i, L_i is the number of persons in the stationary population between ages i and i+1, and $p_{i+0.5}$ is the conditional probability of surviving between age i+0.5 and age i+1.5.

$Var(S_i)$ is estimated by $\sigma^2_{\hat{p}_{i+0.5}}$.

From Chiang (1984, p. 153):

$$\sigma^2_{\hat{p}_i} = \frac{\hat{q}_i^2 (1 - \hat{q}_i)}{D_i}$$

where D_i is the observed number of deaths between ages i and i+1. Chiang derived equation 9 by assuming that D_i is a binomial random variable in N_i trials with probability of dying q_i (Chiang, p. 79). That is, a random number of deaths are assumed to occur to the people alive at age i. This random number of deaths has expected value $N_i q_i$ and variance $N_i q_i(1-q_i)$ (ibid.). Following Chiang, we assume that the number of "trials" N_i is unknown in a current population and estimate $N_i q_i$ by the observed number of deaths D_i.

Since $$\sigma^2_{\hat{p}_i} = \frac{\hat{q}_i^2(1-\hat{q}_i)}{D_i}$$

then by analogy:

$$\sigma^2_{\hat{p}_{i+0.5}} = \frac{\hat{q}_{i+0.5}^2(1-\hat{q}_{i+0.5})}{D_{i+0.5}} \tag{10}$$

Next, we show how equation (10) is calculated in terms of exact ages since life table parameters (eg., q_i) are indexed by exact ages rather than by midyear ages (like $q_{i+0.5}$). First, we derive $q_{i+0.5}$ in terms of life table parameters. By definition $q_{i+0.5} = d_{i+0.5}/l_{i+0.5}$ and $l_{i+0.5} = 0.5(l_i + l_{i+1})$ which is L_i. If deaths are assumed to be uniformly distributed within a year, then $d_{i+0.5} = 0.5(d_i + d_{i+1})$. By substitution,

$$q_{i+0.5} = q_i\left(\frac{.5l_i}{L_i}\right) + q_{i+1}\left(\frac{.5l_{i+1}}{L_i}\right)$$

$$= \frac{d_i}{l_i}\left(\frac{.5l_i}{L_i}\right) + \frac{d_{i+1}}{l_{i+1}}\left(\frac{.5l_{i+1}}{L_i}\right)$$

$$= \frac{.5(d_i + d_{i+1})}{L_i}$$

Then we estimate $D_{i+0.5}$ by $(D_i+D_{i+1})/2$, although it could be estimated by other techniques such as curve fitting.

Finally, equation (10) is restated in terms of exact ages:

$$\sigma^2_{\hat{p}_{i+0.5}} = \frac{\left(\frac{.5(d_i + d_{i+1})}{L_i}\right)^2\left(1 - \frac{.5(d_i + d_{i+1})}{L_i}\right)}{.5(D_i + D_{i+1})} \tag{11}$$

Confidence intervals for the number of net migrants in age-gender group i at time 0 who migrate between time 0 and time 1 are derived from equation (7) with $\sigma_{p_{i+0.5}}$ estimated by the square root of equation (11).

The confidence interval for the total number of net migrants between time 0 and time 1 including all age-gender groups is:

$$\sum_i N\hat{M}F_i(0,1) - 1.96\sigma_{\sum_i N\hat{M}F_i(0,1)} \langle \sum_i NMF_i(0,1) \langle \sum_i N\hat{M}F_i(0,1) + 1.96\sigma_{\sum_i N\hat{M}F_i(0,1)}$$

The variance for the total number of net migrants between time 0 and time 1, summed over all age groups, Var(Σ $NMF_i(0,1)$) is Σ Var $NMF_i(0,1)$ because it is the sum of independent random variables. Total number of net migrants is the sum of independent random variables, the number of net migrants in each age-gender group. These groups are independent because the death rates of any two nonoverlapping age intervals are independent.

The confidence interval for total number of net migrants size cannot be obtained by merely summing the confidence intervals for the age-gender groups. The confidence interval for the net migrants in an age group is a function of the square root of its variance (equation (11)). Although the variance for total net migrants is the sum of the variances for the age groups, the square root of this sum does not equal the sum of the respective square roots.

The preceding discussion presented only the case of estimating net migrants for censuses that are one year apart and assumed that all age groups were close-ended. Next, we consider the more general situation where censuses are t years apart and include a terminal open-ended oldest age group. The following holds for all ages younger than the terminal age-group.

$$\begin{aligned} Var(NMF_i(0,t)) &= Var\left(P_{i+t}(t) - P_i(0)\cdot(S_i \cdot S_{i+1} \cdots S_{i+(t-2)} \cdot S_{i+(t-1)})\right) \\ &= (P_i(0))^2 \cdot Var\left(S_i \cdot S_{i+1} \cdots S_{i+(t-2)} \cdot S_{i+(t-1)}\right) \\ &= (P_i(0))^2 \cdot Var\left(\frac{L_{i+1}}{L_i} \cdot \frac{L_{i+2}}{L_{i+1}} \cdots \frac{L_{i+(t-1)}}{L_{i+(t-2)}} \cdot \frac{L_{i+t}}{L_{i+(t-1)}}\right) \\ &= (P_i(0))^2 \cdot Var\left(\frac{L_{i+t}}{L_i}\right) \\ &= (P_i(0))^2\, Var\left(\frac{.5(l_{i+t} + l_{i+(t+1)})}{.5(l_i + l_{i+1})}\right) \end{aligned}$$

Here the confidence interval takes the following form:

$$\left(NMF_i(0,t) - 1.96\sigma_{N\hat{M}F_i(0,t)}\right)\langle NMF_i(0,t)\langle\left(N\hat{M}F_i(0,t) + 1.96\sigma_{N\hat{M}F_i(0,t)}\right) \qquad (12)$$

Next, we derive the variance of $NMF_i(0,t)$ when the interval is t.

Chiang (1984, p. 156-7) derives the following formula for the sample variance of the probability of surviving from age f to age g:

$$\sigma^2_{\hat{p}_{fg}} = \hat{p}^2_{fg} \cdot \sum_{h=f}^{g-1} \hat{p}_h^{-2} \cdot \sigma^2_{\hat{p}_h} \qquad (13)$$

So substituting equation (13) into equation (8):

$$Var\left(NMF_i(0,t)\right) = \left(P_i(0)\right)^2 \cdot \hat{p}^2_{(i+0.5)(i+t.5)} \cdot \left(\sum_{h=i+0.5}^{i+t.5} \hat{p}_h^{-2} \cdot \sigma^2_{\hat{p}_h}\right) \qquad (14)$$

This leads to the following in terms of life table parameters indexed by exact ages:

$$Var(NMF_i(0,t)) = (P_i(0))^2 (\frac{L_{i+t}}{L_i})^2 \cdot (\sum_{h=i}^{i+(t-1)} \frac{L_{h+1}}{L_h}^{-2} \cdot (\frac{1}{.5(D_h + D_{h+1})}) \cdot \frac{(.5(d_h + d_{h+1})}{L_h})^2 (1 - \frac{.5(d_h + d_{h+1})}{L_h}) \qquad (15)$$

where $\sigma^2_{\hat{p}_h}$ is restated in terms of exact ages according to equation (11).

The final formula requires several alterations to be used for a terminal open-ended age group. One set of changes stems from the fact that, at any time after time 0 the survivors to the open ended age group could come from several (t+1) of the original age groups. For instance, at time t=2 members of the open ended age group 84+ could come from the following age groups at time 0: open ended age group (84+), age 83, or age 82. So the variance of the number of survivors is the weighted sum of the variances for each group that becomes a component of the oldest age group.

$$Var\left(NMF_{84+}(0,t)\right) = \sum_{i=84-t}^{84+} (P_i(0))^2 \cdot \left(\frac{L_{i+t}}{L_i}\right)^2 \cdot \sum_{h=i}^{i+(t-1)} \frac{L_{h+1}}{L_h}^{-2} \left(\frac{1}{.5(D_h + D_{h+1})}\right) \cdot \left(\frac{.5(d_h + d_{h+1})}{L_h}\right)^2 \cdot \left(1 - \frac{.5(D_g + d_h)}{L_h}\right) \qquad (16)$$

Another difference concerns estimation of the survivorship ratio for these ages. Survivorship ratios for the oldest age group are based on T_X, the total number of person-years that would be lived after age x by the cohort of 100,000 births assumed, rather than L_x, the number of person-years lived between ages x and (x+1) by the cohort of 100,000 births assumed. Note that for the last age groups $p_{i+3.5}$ cannot be estimated merely from L_{i+4}/L_i because L for the oldest age group in the life table is too large. So p would be outside the (0,1) interval if it were estimated this way. Instead, survivorship in the last interval should be calculated as it is for survivorship ratios, by the ratio of $T_\omega/T_{\omega-1}$. Also, p cannot be estimated directly but must be viewed in terms of the individual yearly probabilities for each year of the projection. So, for instance, to calculate the confidence intervals for age 81 in a projection from 0 to t=2 using equation (16):

$$
\begin{aligned}
p_{81.5,85.5} &= \frac{l_{85.5}}{l_{81.5}} \\
&= \frac{l_{82.5}}{l_{81.5}} \cdot \frac{l_{83.5}}{l_{82.5}} \cdot \frac{l_{84.5}}{l_{83.5}} \cdot \frac{l_{85.5}}{l_{84.5}} \\
&= \frac{L_{82}}{L_{81}} \cdot \frac{L_{83}}{L_{82}} \cdot \frac{L_{84}}{L_{83}} \cdot \frac{L_{85}}{L_{84}} \\
&= \left(\frac{L_{82}}{L_{81}} \cdot \frac{L_{83}}{L_{82}} \cdot \frac{L_{84}}{L_{83}} \cdot \frac{T_{85}}{T_{84}} \right)
\end{aligned}
$$

Note that for the terminal open-ended age group:

$$
p_{\omega+0.5,\omega+4.5} = \frac{T_{85}}{T_{84}} \cdot \frac{T_{85}}{T_{84}} \cdot \frac{T_{85}}{T_{84}} \cdot \frac{T_{85}}{T_{84}} \tag{17}
$$

To derive confidence intervals for the oldest age group at t=2: using equation (16):

$$
\begin{aligned}
Var(P_{84+}(2)) = {} & P_{84+}(0)^2 \left(\frac{T_{85}}{T_{84}} \cdot \frac{T_{85}}{T_{84}}\right)^2 \cdot \left(\left(\frac{T_{85}}{T_{84}}\right)^{-2} \frac{1}{.5(D_{84}+D_{85})} \cdot \left(\frac{.5(d_{84}+d_{85})}{L_{84}}\right)^2 \left(1 - \frac{.5(d_{84}+d_{85})}{L_{84}}\right) \right) \\
& + P_{83}(0)^2 \cdot \left(\frac{L_{84}}{L_{83}} \cdot \frac{T_{85}}{T_{84}} \right)^2 \cdot \left(\left(\frac{L_{84}}{L_{83}}\right)^{-2} \cdot \frac{1}{.5(D_{83}+D_{84})} \cdot \left(\frac{.5(d_{83}+d_{84})}{L_{84}}\right)^2 \left(1 - \frac{.5(d_{83}+d_{84})}{L_{83}}\right) + \left(\frac{T_{85}}{T_{84}}\right)^{-2} \left(\frac{1}{.5(D_{84}+D_{85})} \left(\frac{.5(d_{84}+d_{85})}{L_{84}}\right)^2 \left(1 - \frac{.5(d_{84}+d_{85})}{L_{84}}\right) \right) \right) \\
& + P_{82}(0)^2 \left(\frac{L_{83}}{L_{82}} \cdot \frac{L_{84}}{L_{83}} \right)^2 \left(\left(\frac{L_{83}}{L_{82}}\right)^{-2} \frac{1}{.5(D_{82}+D_{83})} \left(\frac{.5(d_{82}+d_{83})}{L_{82}}\right)^2 \left(1 - \frac{.5(d_{82}+d_{83})}{L_{82}}\right) + \left(\frac{L_{84}}{L_{83}}\right)^{-2} \frac{1}{.5(D_{83}+D_{84})} \left(\frac{.5(d_{83}+d_{84})}{L_{84}}\right)^2 \left(1 - \frac{.5(d_{83}+d_{84})}{L_{83}}\right) \right)
\end{aligned}
$$

Reverse Survival Method

This procedure estimates the number of net migrants who are age i at time 0, who migrate between time 0 and time t by comparing the actual population age i at time 0 to the population age i+t age time t backcast to time 0.

$$NMR_i(0,t) = \frac{P_{i+t}(t)}{S_i} - P_i(0) \tag{18}$$

The backcasted population is obtained, as presented above, by dividing the number of persons age i+t at time t by S_i, the survivorship ratio (the probability of surviving from age i+.5 to age i+(t+0.5). This procedure assumes that the current life table holds in the past and that S_i does not equal zero. Its variance follows.

$$Var\left(NMR_i(0,t)\right) = Var\left(\frac{P_{i+t}(t)}{S_i} - P_i(0)\right)$$
$$= P_{i+t}(t)^2 \cdot Var\left(\frac{1}{S_{i+t}}\right)$$

Next we derive the variance for the inverse of the survivorship ratio by approximating the inverse by the sum of a convergent power series. Bers (1969, p. 536) shows that the function f(x)=1/x is analytic near every point $x_0 \neq 0$, such that:

$$\frac{1}{x} = \frac{1}{(x-x_0)+x_0}$$
$$= \frac{1}{x_0}\left(1 - \frac{x-x_0}{x_0} + \frac{x-x_0}{x_0} - \ldots\right)$$
$$= \frac{1}{x_0} \cdot \sum_{j=0}^{\infty} \left(\frac{x-x_0}{-x_0}\right)^j.$$

By definition $Var(1/S) = E[1/S^2] - (E[1/S])^2$. We obtain approximations for these expected values by substituting the first few terms of the power series for 1/S. First we approximate 1/S by substituting S for x and $\mu = E[S]$ for x_0 in the above series:

$$\frac{1}{S} = \frac{1}{\mu}\left(1 - \frac{S-\mu}{\mu} + \left(\frac{S-\mu}{\mu}\right)^2 - \ldots\right)$$

Davies (1961) has pointed out that dropping the higher order terms in the power series is a good approximation if the coefficient of variation is small. All applications presented here meet Davies' requirement that the standard deviation should not exceed one fifth of the mean.

Then the expected values of the terms needed for the variance are:

$$E\left[\frac{1}{S}\right] \approx \frac{1+\frac{\sigma^2}{\mu^2}}{\mu}$$

$$E\left[\frac{1}{S^2}\right] \approx \frac{1+3\frac{\sigma^2}{\mu^2}}{\mu^2}$$

where σ^2 is Var(S).

So,

$$Var\left(\frac{1}{S}\right) \approx \frac{1+3\frac{\sigma^2}{\mu^2}}{\mu^2} - \left(\frac{1+\frac{\sigma^2}{\mu^2}}{\mu}\right)^2$$

$$\approx \frac{\sigma^2}{\mu^4} - \frac{\sigma^4}{\mu^6}$$

Substituting this into equation (13) leads to the sample variance of the inverse of the probability of surviving from age f to age g:

$$\sigma^2_{\frac{1}{\hat{p}_{fg}}} \approx \frac{1}{\hat{p}^4_{fg}} \cdot \left(\hat{p}^2_{fg} \cdot \sum_{h=f}^{g-1} \hat{p}_h^{-2} \cdot \sigma^2_{\hat{p}_h}\right) - \frac{1}{\hat{p}^6_{fg}} \cdot \left(\hat{p}^2_{fg} \sum_{h=f}^{g-1} p_h^{-2} \cdot \sigma^2_{\hat{p}_h}\right)^2$$

$$\approx \left[\frac{1}{\hat{p}^2_{fg}} \cdot \sum_{h=f}^{g-1} \hat{p}_h^{-2} \cdot \sigma^2_{\hat{p}_h}\right] - \frac{1}{\hat{p}^2_{fg}} \cdot \left[\sum_{h=f}^{g-1} \hat{p}_h^{-2} \cdot \sigma^2_{\hat{p}_h}\right]^2$$

$$\approx \left[\frac{1}{\hat{p}^2_{fg}} \cdot \sum_{h=f}^{g-1} \hat{p}_h^{-2} \cdot \sigma^2_{\hat{p}_h}\right] \cdot \left(1 - \sum_{h=f}^{g-1} \hat{p}_h^{-2} \cdot \sigma^2_{\hat{p}_h}\right)$$

The variance of the estimate of intercensal net migration from the reverse survival method can be approximated by:

$$Var\left(NMR_i(0,t)\right) \approx P_{i+t}^2(t)\cdot\left(\frac{1}{\hat{p}^2_{(i+0.5)((i+t.5)-1)}}\cdot\sum_{h=i+0.5}^{((i+t.5)-1)}\hat{p}_h^{-2}\cdot\sigma^2_{\hat{p}_h}\right)\cdot\left(1-\sum_{h=i+0.5}^{((i+t.5)-1)}\hat{p}_h^{-2}\cdot\sigma^2_{\hat{p}_h}\right) \quad (19)$$

This leads to the following in terms of life table parameters indexed by exact ages:

$$Var\left(NMR_i(0,t)\right) \approx P_{i+t}^2(t)\cdot\left(\frac{L_{i+t}}{L_i}\right)^{-2}\cdot\left(\sum_{h=i}^{i+(t-1)}\left(\frac{L_{h+1}}{L_h}\right)^{-2}\cdot\frac{1}{.5(D_h+D_{h+1})}\cdot\frac{.5(d_h+d_{h+1})^2}{L_h}\cdot\left(1-\frac{.5(d_h+d_{h+1})}{L_h}\right)\right)$$
$$\left(1-\sum_{h=i}^{i+(t-1)}\left(\frac{L_{h+1}}{L_h}\right)^{-2}\cdot\frac{1}{.5(D_h+D_{h+1})}\cdot\left(\frac{.5(d_h+d_{h+1})}{L_h}\right)^2\cdot\left(1-\frac{.5(d_h+d_{h+1})}{L_h}\right)\right) \quad (20)$$

Average of Forward and Reverse Survival Methods

The final method of estimating the number of net migrants age i at time 0 who migrated between time 0 and time t averages results from the Forward and Reverse Survival Methods.

$$NMA_i(0,t) = .5\left(NMF_i(0,t) + NMR_i(0,t)\right)$$

This section derives confidence intervals from $Var(NMA_i(0,t)$. Both $NMF_i(0,t)$ and $NMR_i(0,t)$ are random variables, so, from the definition of the variance of a linear function of two random variables (Chiang, 1968, p. 15), the variance of their average is:

$$Var\left(NMA_i(0,t)\right) = .5^2\left(Var(NMF_i(0,t) + Var(NMR_i(0,t) + 2Cov(NMF_i(0,t), NMR_i(0,t))\right)$$

where $Cov(NMF_i(0,t), NMR_i(0,t))$ is the covariance of the estimates from the Forward and Reverse Survival methods.

Since the variances of the forward and reverse survival methods were already presented in equations (14) and (19), we next derive the covariance. Recall that the definition of the covariance is Cov(X,Y) = E[XY] - E[X]·E[Y] where X and Y are random variables. Note that Cov(aX+b,cY+d) = ac · Cov(X,Y) (Bickel, and Doksum, 1977).

$$\begin{aligned} Cov(NMF_i(0,t), NMR_i(0,t)) &= Cov\left(P_i(t) - P_i(0)\cdot S_i, \frac{P_i(t)}{S_i} - P_i(0) \right) \\ &= -P_i(0)\cdot P_i(t)\left(Cov(S_i, \frac{1}{S_i} \right) \\ &= -P_i(0)\cdot P_i(t)\left(E\left[S_i \cdot \frac{1}{S_i} \right] - E[S_i]\cdot E\left[\frac{1}{S_i} \right] \right) \\ &= -P_i(0)\cdot P_i(t)\left(1 - E[S_i]\cdot E\left[\frac{1}{S_i} \right] \right) \end{aligned}$$

When we derived the variance of the reverse survival method, we approximated $E[1/S_i]$ using the sum of the convergent power series by

$$E[1/S] \approx (1+\sigma^2/\mu^2)/\mu.$$

So,

$$Cov(NMF_i(0,t), NMR_i(0,t)) \approx -P_i(0)\cdot P_i(t)\left(1 - (\mu \cdot \frac{1+\frac{\sigma^2}{\mu^2}}{\mu}) \right)$$

$$\approx -P_i(0)\cdot P_i(t)\cdot \left(1 - (1 + \frac{\sigma^2}{\mu^2}) \right)$$

$$\approx -P_i(0)\cdot P_i(t)\left(-\frac{\sigma^2}{\mu^2} \right)$$

$$\approx P_i(0)\cdot P_i(t)\left(\frac{\sigma^2}{\mu^2} \right)$$

Next we substitute sample estimators for μ and σ^2.

$$\begin{aligned} Cov(NMF_i(0,t), NMR_i(0,t)) &\approx P_i(0)\cdot P_i(t)(\frac{1}{\hat{p}_{i+0.5}})\hat{p}^2_{(i+0.5)(i+t.5)} \cdot \sum_{h=i+0.5}^{i+t.5} \hat{p}_h^{-2} \cdot \sigma^2_{\hat{p}_h} \\ &\approx P_i(0)\cdot P_i(t) \sum_{h=i+0.5}^{i+t.5} \hat{p}_h^{-2} \cdot \sigma^2_{\hat{p}_h} \end{aligned}$$

Finally, we obtain the variance of the average method by substituting in equations (14) and (19) and the preceding into the second equation in this section:

$$
\begin{aligned}
Var(NMA_i(0,t)) \approx .5^2 \Bigg[& P_i(0)^2 \cdot \hat{p}^2_{(i+0.5)(i+t.5)} \cdot \left(\sum_{h=i+0.5}^{i+t.5} \hat{p}_h^{-2} \cdot \sigma^2_{\hat{p}_h} \right) \Bigg] \\
& + P^2_{i+t}(t) \cdot \left(\frac{1}{\hat{p}^2_{(i+0.5)(i+t.5)}} \cdot \sum_{h=i+0.5}^{i+t.5} \hat{p}_h^{-2} \cdot \sigma^2_{\hat{p}_h} \right) \left(1 - \sum_{h=i+0.5}^{i+t.5} \hat{p}_h^{-2} \cdot \sigma^2_{\hat{p}_h} \right) \\
& + 2 P_i(0) \cdot P_i(t) \cdot \left(\sum_{h=i+0.5}^{i+t.5} \hat{p}_h^{-2} \cdot \sigma^2_{\hat{p}_h} \right)
\end{aligned}
\qquad (21)
$$

Appendix B - Derivation of Adjustment Factors

This section derives the factors for adjusting the variance of the survival ratios in three situations: (1) differences in population size, (2) life table parameters and (3) differences in both size and survivorship.

Adjusting for Differences in Population Size

First, consider the effects of population size on the variance of the survival ratio. We want to estimate intercensal net migration for city A using the life table for state B (A is smaller than B but shares the same life table parameters) and population counts for A. For now, let us assume that we also have information on deaths by age for city A and state B.

Recall the formula for the variance of the probability of dying between ages i and i+1 (equation (11). We want to obtain a formula for the variance of the probability of dying in city A that is a function of the variance for state B and a factor that adjusts for the difference in population sizes between city A and state B. Here is the formula for the variance of the probability dying in state B (the age subscript is assumed but not presented):

$$\sigma^2_{\hat{q}_B} = \frac{1}{D_B} \hat{q}_B \cdot (1 - \hat{q}_B)$$

Let the probability of dying in city A equal the probability of dying in state B for the respective age group. Then,

$$\sigma^2_{\hat{q}_A} = \frac{1}{D_A} \hat{q}^2_A (1 - \hat{q}_A)$$

$$= \frac{1}{D_A} \hat{q}^2_B \cdot (1 - \hat{q}_B)$$

$$= \frac{1}{D_A} \cdot \left(\frac{D_B}{D_B}\right) \hat{q}^2_B \cdot (1 - \hat{q}_B)$$

$$= \frac{D_B}{D_B} \frac{\hat{q}^2_B \cdot (1 - \hat{q}_B)}{D_B}$$

So,

$$\sigma^2_{\hat{q}_A} = \frac{D_B}{D_A} \sigma^2_{\hat{q}_B} \qquad (22)$$

Thus, the confidence interval length is inversely proportional to the square root of the ratio of the number of deaths in state B to the number in city A, which is intuitively appealing. So, for instance, if the number of deaths in city A is 5% of the deaths in state B, the variance is multiplied by 20 and the confidence interval is 4.47 times wider. This example demonstrates how uncertainty increases as the number of deaths becomes small.

This adjustment refers to equations (11) and (13), the sample variance of the survival probabilities. The formulas for the variance of NMF, NMR, and NMA (equations (14), (19) and (21)) are altered only by replacing $\sigma^2_{\hat{p}_h}$ by $D_h / D'_h \; \sigma^2_{ph}$, where $\sigma^2_{\hat{p}_h}$ refers to the variance of the life table survival probability for age h, and D_h/D'_h is the ratio of actual deaths age h in the population with the life table D_h to actual deaths age h for the population in question D'_h.

Sometimes there are applications where the number of deaths in city A for each age-gender group are unknown or unavailable. Then they can be estimated from the population size for city A and the known age-specific death rates for state B (which can readily be obtained from the probabilities of dying in an age-gender group).

Note that an alternative would be to multiply the age-specific mortality rates underlying the "borrowed" life table by the number in each corresponding age group of the population in question. This would preserve the life table survivorship values, but at the same time give the number of deaths appropriate for the population in question. The confidence intervals could then be constructed directly from these "adjusted" deaths.

A note of caution in regard to "size" adjustments is to be certain that differences in deaths are really a reflection of size differences rather than survivorship differences. If questions about this are present, then the preceding alternative may be preferred.

Adjusting for Differences in Survivorship Probabilities

Now we derive an adjustment factor for a situation where we assumed that the number of deaths are equal but the q_x values differ. An example of this usage is when estimating net migration for a city C which has the same number of deaths as city D. The life table for city C is assumed to be some multiple of that for city D.

We assume $D_C = D_D$. We derive the variance of the probability of dying in city C from that for city D and an adjustment factor.

$$\sigma^2_{\hat{q}_C} = \frac{1}{D_C}\hat{q}^2_C \cdot (1-\hat{q}_C)$$
$$= \frac{1}{D_D}\hat{q}^2_C \cdot (1-\hat{q}_C)$$

By cross-multiplying,

$$D_D = \frac{\hat{q}_C \cdot (1-\hat{q}_C)}{\sigma^2_{\hat{q}_C}}$$

So substituting for D_D leads to

$$\sigma^2_{\hat{q}_C} = \frac{\hat{q}^2_C \cdot (1-\hat{q}_C)}{\sigma^2_{\hat{q}_C}} \cdot \frac{\hat{q}^2_D(1-\hat{q}_D)}{\sigma^2_{\hat{q}_D}}$$
$$= \sigma^2_{\hat{q}_D} \cdot \frac{\hat{q}^2_C(1-\hat{q}_C)}{\hat{q}^2_D(1-\hat{q}_D)}$$

In practice where the exact ratio at any given age is not known,we assume a constant ratio across all age groups. This can be approximated, for example, by finding the mean q_x value for the borrowed life table (ignoring q_x for the terminal, open-ended age group) and subjectively estimating the (unknown) mean q_x value for the population in question. A natural approach would be to assume a constant percent difference by age.

We illustrate this procedure with the data from Alaska, 1970-1980. We find that the mean q_x value for the state population with the life table is, 0.034. We then estimate that the desired mean is only 80% of this level in the population in question. Then 0.034*0.8 =0.0275 is the estimated mean q_x values for the Fairbanks population in question. We may then proceed with the adjustment as follows:

$$\sigma^2_{\hat{q}_C} = \sigma^2_{\hat{q}_D} \frac{\hat{q}^2_C(1-\hat{q}_C)}{\hat{q}^2_D(1-\hat{q}_D)}$$
$$= \sigma^2_{\hat{q}_D} \cdot K$$

where K is
$$\sigma^2_{\hat{q}_D} \frac{\hat{q}_C^2(1-\hat{q}_C)}{\hat{q}_D^2(1-\hat{q}_D)}$$

In the example, we have

$$\sigma^2_{\hat{q}_C} \approx \sigma^2_{\hat{q}_D} \cdot \left(\frac{.034^2(1-.034)}{.0275^2(1-.0275)} \right)$$
$$\approx \sigma^2_{\hat{q}_D} \cdot 1.52$$

The actual calculation of the adjusted confidence intervals may be accomplished by first calculating the unadjusted intervals and then multiplying each unadjusted interval by the square root of the adjustment factor. In the example just given, where the adjustment factor is 1.52, this would result in 1.232. Since confidence intervals are a function of the square root of the variance, the adjusted intervals are unadjusted intervals times the square root of the adjustment factor. For instance, in Table 4, the adjusted intervals are 1.232 times the unadjusted intervals since 1.232 is the square root of 1.52.

Adjusting for both Differences in Size and Survivorship: A Short-Cut Method

In a given applicaton, it may be the case that the number of deaths underlying the borrowed life table differ from those in the population in question because of both size and survivorship differences. A quick way to adjust for both types of difference is to use the alternative size adjustment procedure and then apply the adjustment for differences in survivorship.

Acknowledgements

We benefited from discussions with Dana Kamerud (Operating Sciences) about the proofs for reverse survival and average methods and from his comments on a previous version of this paper. Steve Boschee (Pacific Lutheran University) assisted with LOTUS spreadsheets.

References

Alaska, State of. 1983. Alaska Population Overview 1982. Juneau, Alaska: Alaska Department of Labor.

Alho, J. and B. Spencer. 1985. Uncertain population forecasting. Journal of the American Statistical Association 80:306-314.

Bers, L. 1969. Calculus. New York, NY: Holt, Rinehart and Winston.

Bickel, P. and K. Doksum. 1977. Mathematical Statistics: Basic Ideas and Selected Topics. San Francisco, CA: Holden-Day.

Chiang, C. 1984. The Life Table and Its Applications. Malabar, FL: Robert E. Krieger Publishing Company.

Chiang, C. 1968. Introduction to Stochastic Processes in Biostatistics. New York, NY: J. Wiley and Sons.

Cohen, J. 1986. Population forecasts and confidence intervals for Sweden: A comparison of model-based and empirical approaches. Demography 23:105-126.

Davies, O. 1961. Statistical Methods in Research and Production. London, UK: Oliver and Boyd.

Davis, W. 1988. Calculation of the variance of population forecasts. Proceedings of the Business and Economic Statistics Section, American Statistical Association: 567-572.

Espenshade, T. and J. Tayman. 1982. Confidence intervals for postcensal state population estimates. Demography 19: 191-210.

Hamilton, C. H. and F. Henderson. 1944. Use of the survival rate method in measuring net migration. Journal of the American Statistical Association, v. 39:197-274.

Hamilton, C. H. 1961. Some problems in method in internal migration research. Population Index, 27:297-307.

Hamilton, C. H. 1965. Practical and mathematical considerations in the formulation and selection of migration rates. Demography 2:429-463.

Hamilton, C. H. 1966. Effect of census errors on the measurement of net migration. Demography, 3:394-415.

Keyfitz, N. 1981. The limits of population forecasting. Population and Development Review, 7:579-593.

Kintner, H. J. and D. Swanson. 1990. Confidence intervals for the projection of closed populations: A case study of General Motors' salaried retirees. GM Research Laboratories Research Publication GMR-7124.

Kmenta, J. 1971. Elements of Econometrics. New York: MacMillan.

Lee, R.D. 1974. Forecasting births in post-transition populations: Stochastic renewal with serially correlated fertilty. Journal of the American Statistical Association, 69:607-627.

Lee, R. D. 1985. Inverse projection and back projection: A critical appraisal and comparative results for England, 1539 to 1871. Population Studies, 39:233-248.

Pittenger, D.P. 1976. Projecting State and Local Populations. Cambridge, MA: Ballinger.

Pittenger, D.P. 1978. The role of judgment, assumptions, techniques and confidence limits in forecasting population. Socio-Economic Planning Sciences, 12:271-276.

Rives, N.W., Jr. 1982. Assessment of a survey approach. pp. 79-96 in E.S. Lee and H. F. Goldsmith (eds.) Population Estimates: Methods for Small Areas. Beverly Hills, CA: Sage Publications.

Ross, Sheldon 1989. Introduction to Probability Models. Fourth Edition. San Diego, CA: Academic Press.

Saboia, J. 1974. Modeling and forecasting population by time series: the Swedish Case. Demography, 11:483-492.

Shyrock, H., J. Siegel and Associates 1976. The Methods and Materials of Demography. Washington, DC: U.S. Government Printing Office.

Siegel, J. and H. Hamilton 1952. Some considerations in the use of the residual method of estimating net migration. Journal of the American Statistical Association, 47:480-500.

Smith, S.F. 1986. Accounting for migration in cohort-component projections of state and local populations. Demography, 22: 127-134.

Smith, S.F. 1988. Stability over time in the distribution of population forecast errors. Demography, 25:461-474.

Stone, L.O. 1967. Evaluating the relative accuracy and significance of net migration estimates. Demography 4:310-330.

Stoto, M.A. 1983. The accuracy of population projections. Journal of the American Statistical Association, 78: 13-20.

Swanson, D. A. 1989. Confidence intervals for postcensal county-equivalent population estimates: A case study for local areas. Survey Methodology, 15:271-280 (English), 281-290 (French).

United Nations. 1957. Demographic Yearbook. New York: United Nations.

United Nations. 1961. Demographic Yearbook. New York: United Nations.

United Nations. 1970. Manual VI: Methods of Measuring Internal Migration. Department of Economic and Social Affairs, Population Studies No. 47. New York: United Nations.

United States, Bureau of the Census. 1961. Census of Population: 1960, Volume 1, Characteristics of the Population, Part 32, New Jersey. Washington, DC: U. S. Government Printing Office.

United States, Bureau of the Census. 1961. Census of Population: 1960, Volume 1, Characteristics of the Population, Part 34, New York. Washington, DC: U. S. Government Printing Office.

United States, Bureau of the Census. 1961. Census of Population: 1960, Volume 1, Characteristics of the Population, Part 40, Pennsylvania. U. S. Government Printing Office.

United States, Bureau of the Census. 1973. Census of Population: 1970, Volume 1, Characteristics of the Population, Part 32, New Jersey. U. S. Government Printing Office.

United States, Bureau of the Census. 1973. Census of Population: 1970, Volume 1, Characteristics of the Population, Part 34, New York. U. S. Government Printing Office.

United States, Bureau of the Census. 1973. Census of Population: 1970, Volume 1, Characteristics of the Population, Part 40, Pennsylvania. U. S. Government Printing Office.

United States, Bureau of the Census. 1961. Census of Population: 1960. General Population Characteristics, PC(1)-B. U. S. Government Printing Office.

United States, Bureau of the Census. 1972. Census of Population: 1970. General Population Characteristics, PC(1)-B. U. S. Government Printing Office.

United States, Bureau of the Census. 1983. Current Population Reports, Series P-25, n. 937. Washington, DC: U. S. Government Printing Office.

United States, Public Health Service. 1963. Vital Statistics of the United States: 1960, Volume II - Mortality Part B. U.S. Department of Health, Education, and Welfare. U.S. Government Printing Office.

United States, Public Health Service. 1974. Vital Statistics of the United States: 1970, Volume II - Mortality Part B. U.S. Department of Health, Education, and Welfare. U.S. Government Printing Office.

United States, Public Health Service. 1964. Vital Statistics of the United States: 1960, Volume I - Fertility. U.S. Department of Health, Education, and Welfare. U.S. Government Printing Office.

United States, Public Health Service. 1965. Vital Statistics of the United States: 1961, Volume I - Fertility. U.S. Department of Health, Education, and Welfare. U.S. Government Printing Office.

United States, Public Health Service. 1966. Vital Statistics of the United States: 1962, Volume I - Fertility. U.S. Department of Health, Education, and Welfare. U.S. Government Printing Office.

United States, Public Health Service. 1967. Vital Statistics of the United States: 1963, Volume I - Fertility. U.S. Department of Health, Education, and Welfare. U.S. Government Printing Office.

United States, Public Health Service. 1968. Vital Statistics of the United States: 1964, Volume I - Fertility. U.S. Department of Health, Education, and Welfare. U.S. Government Printing Office.

United States, Public Health Service. 1969. Vital Statistics of the United States: 1965, Volume I - Fertility. U.S. Department of Health, Education, and Welfare. U.S. Government Printing Office.

United States, Public Health Service. 1970. Vital Statistics of the United States: 1966, Volume I - Fertility. U.S. Department of Health, Education, and Welfare. U.S. Government Printing Office.

United States, Public Health Service. 1971. Vital Statistics of the United States: 1967, Volume I - Fertility. U.S. Department of Health, Education, and Welfare. U.S. Government Printing Office.

United States, Public Health Service. 1972. Vital Statistics of the United States: 1968, Volume I - Fertility. U.S. Department of Health, Education, and Welfare. U.S. Government Printing Office.

United States, Public Health Service. 1973. Vital Statistics of the United States: 1969, Volume I - Fertility. U.S. Department of Health, Education, and Welfare. U.S. Government Printing Office.

United States, Public Health Service. 1974. Vital Statistics of the United States: 1970, Volume I - Fertility. U.S. Department of Health, Education, and Welfare. U.S. Government Printing Office.

Voss, Paul, C. Palit, B. Kale, and H. Krebs. 1981. Forecasting State Populations Using ARIMA Time Series Techniques. Madison: University of Wisconsin, Applied Population Laboratory.

Received, December 1992; revised September 1993.