On Future Population
Author(s): Nathan Keyfitz
Source: *Journal of the American Statistical Association*, Vol. 67, No. 338 (Jun., 1972), pp. 347-363
Published by: Taylor & Francis, Ltd. on behalf of the American Statistical Association
Stable URL: https://www.jstor.org/stable/2284381
Accessed: 01-06-2020 07:31 UTC

© Journal of the American Statistical Association
June 1972, Volume 67, Number 338
Invited Papers Section

On Future Population

NATHAN KEYFITZ*

Population projections as officially published show the consequences of a set of assumptions on births, deaths and migration. Population forecasts are unconditional statements of what the population of a given area at some future date will be, preferably in the form of a probability distribution. Methods for projections, and the problems in applying them to forecasting, are described. Demographic forecasting is seen as the search for functions of population that are constant through time, or about which fluctuations are random and small. The possibilities for such functions are described as well as the results of their application to United States and world populations.

1. INTRODUCTION

The future is very real to those who study population. The question that they typically address to data is, "What would happen if these rates continued?," a question in which much of demographic technique originates. If the internal demands of his discipline ever let the demographer forget the future, his clients bring him back to it, with pleas for more and better population forecasts.

The future populations calculated by demographers are typically *projections*, the numerical consequences of the assumptions chosen. The numbers are conditional on the assumptions being fulfilled: *if* birth and death rates move in a certain fashion, and net migration is of such and such an amount, then the total population and its age-sex distribution will be such and such. They are correct beyond any testing against the subsequent population performance; in fact they can be incorrect only in the trivial sense that the author made an arithmetic error that prevents his final numbers from being consistent with his initial assumptions.

But the user is typically not a professional demographer and he wants a *prediction* of what will actually happen in the future. The calculation of future population, presented as an innocent (indeed tautological) projection by its author, is accepted as a prediction by the reader. The bridge between the two points of view is in the choice of the assumptions. Insofar as these are realistic the projection does indeed forecast what will happen. Hence the stress that will be laid in what follows not only on the nature of the assumptions made, but on their presentation in such a way that the user is genuinely in a position to decide on their merits.

* Nathan Keyfitz is Andelot Professor of Demography and Sociology, Harvard University, Cambridge, Mass. 02138. He is grateful to Byron W. Brown, John Durand, Morris L. Eaton, M. A. El-Badry, Gus Haggstrom, Jan Hoem, William Kruskal, Robert Lundy, John McDonald, Paul Meier, Lincoln Moses, Samuel H. Preston, Andrei Rogers, Richard J. Savage, Robert Schmitt and Jacob S. Siegel for their comments. Financial support was provided by NSF grant GZ995, and NIH research contract 69-2200.

First a review of the mechanics of population projection. The reader already acquainted with these, or interested in broader aspects of the prediction problem, can skip this review and go on to Section 3.

2. PROJECTION

Even crude birth and death rates may be illuminated by asking what would happen if they were to continue. If the number of births in a recent year was B, the number of deaths was D, and the population at mid-year P, our first step in attempting to make sense of these figures is to calculate crude rates of birth and of death, B/P and D/P respectively, and then take the difference between them, $B/P-D/P$, the crude rate of natural increase. This was 0.035 for Mexico and 0.008 for the United States, both in 1969. We can say that Mexico is increasing at $3\frac{1}{2}$ percent per year, the United States at 0.8 percent per year. In any particular year the difference in the effect on the total is slight, but if the rates continue for 100 years Mexico will grow 32 times, while the United States somewhat more than doubles. Projection is like a microscope that magnifies the differences of a given period, and so helps analysis and understanding of current rates.

Projection methods are distinguished by what they hold constant into the future. In the preceding paragraph we supposed the crude rate of natural increase, $B/P-D/P$, to be held constant. We could alternatively suppose that the difference between the crude rate of last year and that of this year would be constant, that is, make a straight line extrapolation of the rate of increase. We could hold constant second, third or higher order differences of the rate, corresponding to extrapolation by a second, third or higher order polynomial. The same primitive extrapolation could be applied directly to the population total itself, avoiding rates; if P_t is the population at time t, then $P_t=P_0+t(P_0-P_{-1})$ if first differences are constant, and this is often used for very short term projections.

When age distributions are irregular because of past fluctuations in the birth rate or other reasons, it seems preferable to extrapolate on age-specific birth and death rates, assuming for example that the last rates observed will continue into the future, or that first, second or other differences of the time series of rates at each separate age will continue.

But this does not take account of the possibility that individual couples may aim at a certain number of children, say three, and differ only in the spacing on which they have them. Where this is so the extrapolation of age-specific rates is less satisfactory than extrapolation of number of children in completed families.

In arranging this part of the exposition according to what features of the past are to be held constant into the future, we start with fixed age-specific rates of death and suppose zero births and zero migration.

2.1 Survivorship in a Cohort Line

Consider a cohort or number of babies born at one time, say B in number. Call the expected number alive a year later $B\ell_1$, the expected number alive five years later $B\ell_5, \cdots$, and x years later $B\ell_x$. If $\ell_0=1$, then $\ell_1, \ell_5, \cdots, \ell_x$ are the probabilities that one individual baby will live for $1, 5, \cdots, x$ years; our model is deterministic—it supposes that the fraction surviving will be exactly equal to the probability of individual survival.

The survivorship argument applies indifferently to cohorts of males, females, or the two sexes together. The cohort consisting, for example, of $B=100{,}000$ girls born at one time, is represented by the ℓ_x column of a life table as usually published, which corresponds to $100{,}000\ell_x$ in our notation.

But in reality a cohort is not born at a moment of time; its births are spread out over a period, commonly taken as five years. Suppose births take place uniformly over a five-year period ending at time $t=0$, at the rate of B per year, so that the births between time t and $t+dt$ are Bdt. Expected survivors to the end of the five-year period from among these will be $B\int_0^5 \ell_t dt$, which may be written $B{}_5L_0$. By the same argument the expected number of survivors at times $x=0, 5, 10, \cdots$ years after the end of the five-year period over which the cohort was born will be $B{}_5L_x$, and at the end of $x+n$ years will be $B{}_5L_{x+n}$, where n as well as x is a multiple of 5, and ${}_5L_x$ is defined as ${}_5L_x=\int_0^5 \ell_{x+t}dt$. (In an increasing population the cohort starts life with an increasing number of births within its five-year period. If the increase is uniform ${}_5L_x$ would be replaced in all that follows by ${}_5L_x' = \int_0^5 e^{-r(t-2.5)}\ell_{x+t}dt$ where r is the rate of natural increase, but the difference between ${}_5L_x$ and ${}_5L_x'$ rarely rises as high as one percent [23, p. 215].

This leads to the population projection for those persons alive at time 0, the jumping-off point. If the number of individuals observed of ages x to $x+4$ at last birthday is ${}_5P_x$, then the number n years later will be $({}_5P_x/{}_5L_x){}_5L_{x+n}$. This supposes no net migration over the n years, that the ratio ${}_5L_{x+n}/{}_5L_x$ comes from a life table appropriate to the group under discussion, and that the survival ratios for cohorts are exactly equal to the probabilities for individuals, as characterizes a deterministic model.

Such a projection follows the prospective path of the population numbers down cohort lines. Given an initial female population, say ${}_5P_x^{(0)}$, we can fill out a triangular table:

$$\begin{array}{ccccc}
{}_5P_0^{(0)} & — & — & — & — \\
{}_5P_5^{(0)} & {}_5P_5^{(1)} & — & — & — \\
{}_5P_{10}^{(0)} & {}_5P_{10}^{(1)} & {}_5P_{10}^{(2)} & — & — \\
{}_5P_{15}^{(0)} & {}_5P_{15}^{(1)} & {}_5P_{15}^{(2)} & {}_5P_{15}^{(3)} & — \\
\vdots & \vdots & \vdots & \vdots & \vdots
\end{array} \tag{2.1}$$

in which the survivors to successive later periods are identified with superscripts showing how many five-year periods have gone by. In general ${}_5P_{x+5t}^{(t)} = {}_5P_x^{(0)}({}_5L_{x+5t}/{}_5L_x)$. Everyone is supposed dead before age ω, a multiple of 5, so that ℓ_ω and ${}_5L_\omega$ are zero. The dashes in the upper right triangle of (2.1) are for the cohorts that would be initiated by births after time zero.

We will find it convenient for discussion, though not necessarily for computation, to express the same projection in the form of a matrix operator. If the column of ${}_5P_x^{(0)}$, the numbers of females at time zero, be written as the vertical array or vector $\boldsymbol{P}^{(0)}$, and if we define a survivorship matrix $\boldsymbol{S}$ as

$$\boldsymbol{S} = \begin{bmatrix}
0 & 0 & 0 & \cdots \\
\dfrac{{}_5L_5}{{}_5L_0} & 0 & 0 & \cdots \\
0 & \dfrac{{}_5L_{10}}{{}_5L_5} & 0 & \cdots \\
\vdots & & & \vdots
\end{bmatrix} \tag{2.2}$$

such that the jth element of the ith row is

$$\begin{aligned}
s_{ij} &= \frac{{}_5L_{5i-5}}{{}_5L_{5i-10}}, \qquad i = j+1 = 2, 3, \cdots, \\
&= 0, \qquad \text{otherwise},
\end{aligned}$$

then the projection up to this point can be written as a recurrence equation, relating the survivors at time t to those at time $t-1$:

$$\begin{aligned}
\boldsymbol{P}^{(1)} &= \boldsymbol{S}\boldsymbol{P}^{(0)} \\
\boldsymbol{P}^{(2)} &= \boldsymbol{S}\boldsymbol{P}^{(1)} \\
&\vdots \\
\boldsymbol{P}^{(t)} &= \boldsymbol{S}\boldsymbol{P}^{(t-1)}.
\end{aligned}$$

The solution of the recurrence set that gives $\boldsymbol{P}^{(t)}$ in terms of $\boldsymbol{P}^{(0)}$ without going through intermediate values is

$$\boldsymbol{P}^{(t)} = \boldsymbol{S}^t\boldsymbol{P}^{(0)}.$$

So much for the survivors of those alive at the start of the projection. We now need the future births to fill out the upper right triangle of the projection (2.1).

2.2. The Fertility Component

Suppose that, among B female births taking place uniformly over an observed five-year period in the past, ${}_5B_x$ occurred to women aged x to $x+4$ at last birthday and that the average number of such women exposed was ${}_5P_x$ in the same age interval. Then the age-specific birth rate was $F_x = {}_5B_x/{}_5P_x$. (Where the prescript 5 is suggested by the context it may be omitted.)

The set of F_x, $x=\alpha, \alpha+5, \cdots, \beta-5$, are given by observation or hypothesis, α and β being multiples of 5 that embrace the ages of childbearing. To find the births in the projection we apply them to the mean survivors for each future time. For example, in the first time period and for the age interval x to $x+4$ we start out with ${}_5P_x^{(0)}$ women and end with ${}_5P_x^{(1)} = {}_5P_{x-5}^{(0)}({}_5L_x/{}_5L_{x-5})$. One approximation to the mean survivors or average exposure per year during the five years is the arithmetic mean of these two numbers, the contribution of women aged x to $x+4$ to the expected births being obtained by multiplying the average exposure by $5F_x$:

$$\left(\frac{{}_5P_x^{(0)} + {}_5P_{x-5}^{(0)}({}_5L_x/{}_5L_{x-5})}{2}\right)5F_x, \qquad x = \alpha, \alpha+5, \cdots, \beta-5. \tag{2.3}$$

The grand total of births would be the sum of such expressions over x from α to $\beta-5$, since the age range α, β includes all positive fertility.

On a fixed life table the births during a five-year time interval determine the population 0–4 at the end of the interval. We multiply the births as estimated in (2.3) by ${}_5L_0/5$ to obtain the survivors to the end of the five-year period, and add through x to obtain ${}_5P_0^{(1)}$:

$${}_5P_0^{(1)} = {}_5L_0 \sum_{x=\alpha}^{\beta-5} \left[{}_5P_x^{(0)} + {}_5P_{x-5}^{(0)}({}_5L_x/{}_5L_{x-5})\right]F_x/2; \tag{2.4}$$

successive application of the procedure exemplified in (2.4) fills out the upper right triangle of (2.1).

This part of the operation may be translated into a matrix $\boldsymbol{B}$ that has zeros everywhere except in its first row. Because we will want $\boldsymbol{B}$ to premultiply the vector $\boldsymbol{P}$ we will have to rearrange the sum on the right-hand side of (2.4). The element ${}_5P_x^{(0)}$, for example, occurs in two terms of (2.4), in one of which it has a coefficient ${}_5L_0F_x/2$, and in the subsequent one of which it has a coefficient ${}_5L_0({}_5L_{x+5}/{}_5L_x)F_{x+5}/2$. Thus the rearranged sum for ${}_5P_{(0)}^{(1)}$ is

$${}_5P_0^{(1)} = \sum_{x=\alpha}^{\beta-5} \left[\frac{{}_5L_0}{2}\left(F_x + \frac{{}_5L_{x+5}}{{}_5L_x}\;F_{x+5}\right)\right]{}_5P_x^{(0)}, \tag{2.5}$$

which is the first element of the product vector of $\boldsymbol{B}$ and $\boldsymbol{P}^{(0)}$. The quantity in square brackets in (2.5) is the $[(x/5)+1]$th element of the first row of $\boldsymbol{B}$, corresponding to age x to $x+4$ at last birthday. All elements of $\boldsymbol{B}$ below the first row are zero.

This completes the projection of a closed population with fixed age-specific rates of birth and death. The age-specific death rates have led to a life table whose ${}_5L_x$ appears in our projection matrix $\boldsymbol{S}$, and the age-specific rates F_x by which women give birth to girl children appear along with the ${}_5L_x$ in $\boldsymbol{B}$. The combined application of survivorship and birth is described by the matrix sum $\boldsymbol{S}+\boldsymbol{B}$, which we will write $\boldsymbol{M}$. Unless projections of the several generations are required separately, or birth and suvivorship are to change at different rates, $\boldsymbol{S}$ and $\boldsymbol{B}$ need not be retained as such. The population at intervals of five years after the initial time 0 is given by the set of equations $\boldsymbol{P}^{(t)} = \boldsymbol{M}\boldsymbol{P}^{(t-1)}$, $t=1, 2, \cdots$. The solution of these equations is

$$\boldsymbol{P}^{(t)} = (\boldsymbol{S}+\boldsymbol{B})^t\boldsymbol{P}^{(0)} = \boldsymbol{M}^t\boldsymbol{P}^{(0)}. \tag{2.6}$$

This simplest formula of population projection supposes fixed age-specific rates of birth and death and zero net migration.

2.3 Projection for Two Sexes

One could carry out the projection (2.6) for females first, and then for males, using for each its own life table and the age-specific rates of motherhood and fatherhood respectively. To deal with the two sexes separately can produce discrepancies because the ratio of increase based on statistics of any one year is in general somewhat different for males and for females.

Though the point is not of great importance for those medium-term projections in which interest is mainly concentrated, nonetheless most writers avoid discrepancies by attributing births of boys as well as of girls to one parent. The age distribution of mothers at the birth of children suffices to project both sexes on what is called a female dominant model, for which ages of fathers are not required. The number of boys born in each projection cycle is taken as equal to the number of girls born multiplied by the sex ratio at birth, that sex ratio being in the neighborhood of 1.05. The remainder of the projection is as before: survivorship of males and females is according to their respective life tables. The ultimate rate of increase is not affected by parameters referring to the population which is either male or past reproduction age.

A matrix that incorporates both sexes may be readily assembled on the female dominant model, and (2.6) continues to hold with $\boldsymbol{M}$ interpreted as this larger matrix, and $\boldsymbol{P}$ as a vector in which the age-sex distribution is contained. If the upper left quarter of $\boldsymbol{M}$ projects females, then the upper half of $\boldsymbol{P}$ contains females by age. By an obvious modification dominance may be male, or mixed.

2.4 History of Projection

In (2.6) we have reached the population projection with fixed rates taking account of age. This modern approach, based on separate recognition of birth and death components, seems first to have been used by Edwin Cannan [10]. He in effect applied our matrix $\boldsymbol{M}$, but took a less exact account of the age distribution of childbearing.

Bowley [7] also followed each sex down its cohort line for survivorship, but used women 20 to 44 years of age as the base for births. The next appearance of the components method, a thoroughgoing exploration that has brought it into very wide use, was by Whelpton [62, 63]. It remained for Bernardelli [4], Lewis [29], and above all Leslie [26, 27, 28] to show that the procedure could be compactly presented in matrix form, and that such a presentation permits analysis of the regime of mortality and fertility in abstraction from the age distribution by which it is to be multiplied. The continuous form of the same projection was imaginatively studied by Lotka [31, 32, 51] over a 40-year career that created the core of theoretical demography.

2.5 Asymptotic Properties of the Linear Projection Model

The projection serves to interpret a particular set of observed birth and death rates, and should describe exactly the operation of these. If the observations come in age intervals of finite width, such as five years, then both the life table and the projection are somewhat indeterminate [21, Ch. 1; 22] because of our ignorance of the distribution of the observed population within each five year age interval, but the effect of this is small.

The method is somewhat biased in supposing that the average population exposed to the risk of child-bearing is the arithmetic mean embodied in (2.3) of the initial and final numbers, age group by age group, within each projection cycle. This form of calculation, chosen for the convenience of its linearity, may be shown [21, p. 255], to exaggerate slightly the rate of increase of the population.

Once the projection process is arranged in the matrix form (2.6) we can bring the theory of matrices with nonnegative elements to the analysis of its properties. These have been studied by Leslie [26], Lopez [30], McFarland [38], Sykes [53], and most recently by Parlett [43].

Except in the practically unimportant case where fertility is confined to a single age group, or to two age groups that are not prime to one another, the matrix $\boldsymbol{M}$ will fill up with positive numbers as it is taken to a high power. Sooner or later $\boldsymbol{M}^{t+1}$ will approximate a simple multiple, say λ, of $\boldsymbol{M}^t$, so we may write $\boldsymbol{M}^{t+1} \doteq \lambda \boldsymbol{M}^t$. The ratio λ by which the matrix increases from one power to the next when t is large is also the ratio in which the population increases under continued projection. Knowledge of ultimate or asymptotic properties is useful: it helps us to understand the set of observed age-specific birth and death rates. The intrinsic rate of natural increase, r, is given by $0.2 \log \lambda$; r is the ultimate rate of increase implied by the regime of mortality and fertility.

The quantity $\boldsymbol{M}^t \boldsymbol{P}^{(0)}/\lambda^t$, for t large, is a vertical vector whose projection into the future agrees asymptotically with the projection of the observed population. It is called the stable equivalent of the observed population. Any multiple of this stable equivalent may be designated $\boldsymbol{K}_1$, a positive vector whose direction (i.e., the ratios of its elements) is fixed under projection. One convenient form, the stable age distribution, is obtained by dividing the number at each age by the total at all ages. From one point of view the whole theory of stable age distributions is a part of the subject of population projection: that part concerned with what happens with the continuance of a set of age-specific birth and death rates into the indefinite future in the absence of migration. But such considerations take us too far from the more immediate issues of projection, to some of whose more concrete problems we now return.

2.6 Homogeneity

Careful work in projection starts by dividing the population into homogeneous subpopulations. In the United States, for example, we know that the birth rate for Blacks is higher than for Whites, not only now but most probably for a considerable time in the future. Deaths are also higher, but their effect on projection is less, and on balance Blacks are increasing faster than Whites. If the two subpopulations are not distinguished, but the whole is projected by average birth and death rates, a systematic *understatement* results, as compared with treating Blacks and Whites separately and projecting each by its own birth and death rates. The proof of this is omitted here, except to say that it depends only on the fact that an arithmetic mean is greater than a geometric mean. Recognizing subpopulations with distinguishable regimes of mortality and fertility permits in the calculation—what inevitably happens in real populations—the group with the more rapid rate of increase to rise as a fraction of the total.

As an example that will suggest the quantitative effect, the United States for 1966 without breakdown, as projected with a life table made from the deaths in that year and the fertility rates implied by the births of that year, results in a total population for 1981 of 230,477,000. Recognizing the two separate groups of Whites and Nonwhites, whose population, births and deaths add exactly to the totals for the United States, constructing life tables and age-specific birth rates for the two groups separately, and then projecting each by means of its own life table and birth rates, gives 1981 Whites as 199,287,000, and Nonwhites as 31,441,000, which add to 230,728,000. The deficiency of the projection not recognizing color is 251,000, or about one part in a thousand, after 15 years. After 100 years the separate projections add to 8 percent more than the projection without breakdown.

2.7 Changing Rates of Birth and Death

The age-specific rates of birth and death embodied in $\boldsymbol{M}$ have been supposed either observed for some past period or hypothetical, perhaps derived from model tables, but in any case fixed as the projection goes through time. To liberalize the projection, we now allow the birth and death rates to change. We may for instance note in a series of past observations that birth and death rates are declining linearly, and instead of assuming fixed age-specific rates, assume fixed first differences of rates.

If the projection matrix for the first interval after the jumping-off point is $\boldsymbol{M}^{(1)}$, that for the second interval is $\boldsymbol{M}^{(2)}, \cdots$, then the population at the end of the tth interval is

$$P^{(t)} = M^{(t)}P^{(t-1)}, \qquad t = 1, 2, \cdots \tag{2.9}$$

and by successive substitution $P^{(t)}$ may be expressed in terms of $P^{(0)}$ as

$$P^{(t)} = M^{(t)}M^{(t-1)} \cdots M^{(1)}P^{(0)}. \tag{2.10}$$

Such projections for large t were studied by Coale and Lopez [30]. No fixed ultimate or stable population exists as for projection with a fixed matrix, but if two populations, of different age distributions, are subjected to the same sequence of regimes of mortality and fertility, that is to the same sequence of $M^{(1)}, M^{(2)}, \cdots$, then their age distributions will become more and more alike. Weak ergodicity was the name given to this property by Coale.

Suppose future decline in the birth rate takes place at p percent per five-year period for each age of mother. Then calling $M^{(0)} = S + B$ the initial projection matrix, typically referring to the immediate past, the projection matrix for the first future period will be $M^{(1)} = S + B[1-(p/100)]$, for the second period will be $M^{(2)} = S + B[1-(p/100)]^2$, etc. Arithmetical implementation of (2.10) with these particular values of $M^{(1)}, M^{(2)}, \cdots$, will provide $P^{(t)}$ for births declining in this particular way.

Matrix expressions may similarly be found to express falling mortality. In practice, once a computer subroutine for a life table is available, it is more convenient to extrapolate the age-specific death rates and have the subroutine make a fresh life table on each projection cycle.

2.8 Extrapolation of Death Rates at Individual Ages

We may suppose that the age-specific death rates follow geometric or arithmetic progression, each age having its own constants, for 10 or 20 years, but the longer the projection the less satisfactory this will be. An alternative is to fit a curve such as $\alpha + \beta\gamma^t$ or $\beta\gamma^{\delta t}$ to the age-specific death rates [36]. With a computer one can apply the least squares criterion to either curve in an iterative process.

To allow each separate five-year age group to establish its own trend for extrapolation will often be too much of a concession to flexibility, and produce erratic results, especially for small populations. On the other hand to suppose that change in mortality occurs uniformly at all ages is to go against all experience. As a minimum we must distinguish three age classes: (a) infancy, (b) childhood, youth and middle age, and (c) old age. Decrease in infant mortality is similar in its population effect to increase in births; changes in mortality at child-bearing ages have more complex effects on later population.

Up to a few years ago the trend of mortality in all countries was downward. Now Western Europe and the United States seem to have levelled off, while most other parts of the world are dropping rapidly. Quite different techniques for projection are suggested for the two circumstances. Where mortality is changing slowly or not at all one magnifies random variation by extrapolating past age-specific rates; the statistical problem is to decide whether the trend of past rates stands out clearly enough from random fluctuations that extrapolation of the trend offers a net gain over supposing that the recent average death rates will continue into the future.

2.9 Order of Birth

The study of birth order is especially important in a time of changing patterns of fertility. The matter came to general attention in the early 1940's, when so many marriages took place, followed in a high proportion of instances by births, that the number of first births in some years was more than the number of individuals of any one cohort entering childbearing ages. Such a piling up of first births, concealed within the total births of certain years, gives an overall rate that in its nature cannot continue for very long. To understand what is happening we need to see what fraction of marriages (considered as births of zero order) progress to first births, what fraction of couples having first births progress to second births, etc. The consideration of parity progression leads naturally to a tracing of individual cohorts, and Akers [2] describes the experience of the Bureau of the Census in this. The pioneer exploration is due to Whelpton [64], and Ryder [49] provides a recent summary.

Recognizing parity achieves more specificity than the usual recognition of age and sex. If we classify births by order as well as age of the mother, and divide by the number of exposed women who are in the age group and have had a child of one lower order, then we obtain age-parity-specific rates. From these rates, probabilities [42] and a projection matrix [41] can also be worked out by means analogous to those for age projection.

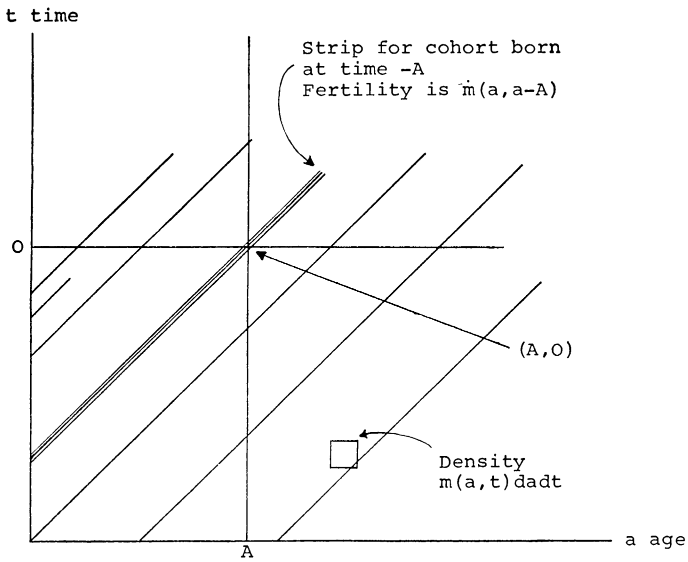
2.10 Cohorts versus Periods

Everything prior to Section 2.9 is in terms of aggregate births of the several years or other periods, considered as a time series. If each family made its decisions on having or not having children in relation to the conditions of the year, without reference to the number of children it has had in the past, nothing more need be said. But suppose now that couples aim at a certain number of children; good or bad times cause them to defer or to advance their childbearing, but not to change the total number. Then the fluctuations in the time series of births are less consequential; the drop of the birth rate in a depression would be made up in the subsequent business upswing, and the rate of increase of the population would be lower only insofar as older parents imply a greater length of generation. Constancy in the total number of children per family results in constancy in the average number of children per cohort, and the Bureau of the Census and the Scripps Foundation have taken advantage of this constancy in their projections.

The trouble with cohorts is that we do not know their completed size until their childbearing is over. If we project from completed cohorts only we handicap ourselves by disregarding information contained in those cohorts that are still bearing, and this omission would cancel out the advantage over the period method, where the latest information is all routinely put to use. Thus the method of cohorts requires some way of translating period into cohort distributions.

The easy way to effect this translation is in terms of moments. Consider the figure, known as a Lexis diagram, in which the life lines of individual women are plotted.

LEXIS DIAGRAM FOR m(a, t) SHOWING SHIFT OF ORIGIN TO POINT (A, 0)



Small squares are used to indicate the ages and times when women bear children. Think of women aged a at time t and suppose that $m(a, t)$ is the age-specific rate of childbearing for those women. Then the cohort of women that is aged A at time 0 would be represented by $m(a, a-A)$ as it passed through the several ages a. Its Gross Reproduction Rate (GRR) would be $R_0^*(A)=\int_\alpha^\beta m(a, a-A)da$, the star to identify $R_0^*(A)$ as referring to a cohort. The nth moment around A would be

$$\frac{R_n^*(A)}{R_0^*(A)}=\frac{\int_\alpha^\beta (a-A)^n m(a, a-A)da}{R_0^*(A)}.$$

Corresponding parameters of the period distribution at time zero are $R_0=\int_\alpha^\beta m(a, 0)da$, the ordinary period GRR, and

$$\frac{R_n(A)}{R_0}=\frac{\int_\alpha^\beta (a-A)^n m(a, 0)da}{R_0},$$

the nth period moment around A. (Note that R_0 does not depend on A.)

Now our problem is to express the cohort $R_n^*(A)$ in terms of the period $R_n(A)$. Using Taylor's theorem to expand $m(a, a-A)$ under the integral sign,

$$\begin{aligned} R_n^*(A) &= \int_\alpha^\beta (a-A)^n m(a, a-A)da \\ &= \int_\alpha^\beta (a-A)^n [m(a, 0) + (a-A)\dot{m}(a, 0) \\ &\qquad + ((a-A)^2/2!)\ddot{m}(a, 0) + \cdots]da \end{aligned}$$

where $\dot{m}(a, 0)$ is the first derivative of $m(a, 0)$ with respect to time, $\ddot{m}(a, 0)$ the second derivative, $\cdots$. Using the facts that under conditions general enough for our purpose (1) the integral of a sum is the sum of the integrals and (2) the integral of a derivative is the derivative of the integral, we obtain the simple and important result

$$R_n^*(A) = R_n(A) + \dot{R}_{n+1}(A) + \frac{\ddot{R}_{n+2}(A)}{2!} + \cdots \quad (2.11)$$

due to Ryder [48]. This serves to estimate any cohort moment about any age A in terms of the period moments and their changes. It applies equally to net reproduction if the $m(a, t)$ is replaced by $\ell_a m(a, t)$.

The most useful special case is that for $n=0$,

$$R_0^*(A) = R_0 + \dot{R}_1(A) = R_0\left(1 + \frac{\dot{R}_1(A)}{R_0}\right), \quad (2.12)$$

which we have truncated as though the second derivatives of the second and higher moments are zero. This result has a simple intuitive meaning. Suppose that the period GRR is unchanging, but that the average age of motherhood is increasing through successive periods. Then $\dot{R}_1(A)$ will be positive, and the cohort $R_0^*(A)$ will be greater than the period R_0. The derivative $\dot{R}_1(A)$ gives the amount by which the period moment is a distorted version of the cohort moment.

Since the result (2.11) applies for any A, and since the distribution of the cohort childbearing is implicit in its moments, we can use this in principle to fill out the future childbearing of the incomplete cohorts at time zero.

If we knew that every cohort was aiming at exactly three children (2.11) and (2.12) would not be needed; we would simply deduct the average number of children already recorded from (2.3), and suppose the remainder distributed over time and age in the future in some suitable way. If on the other hand cohorts had nothing to do with the matter we would treat the births to women of given age in successive periods as an ordinary time series and extrapolate. The theory is especially useful for the intermediate case where the R_0^* for cohorts are shifting less rapidly than the R_0 for periods. To take advantage of this we extrapolate $R_0^*(A)$ and one or two further moments, then fill in the cells of the table for future tides and ages, for example by supposing a normal or gamma distribution for childbearing in each cohort. This seems to provide a better estimate of childbearing in each future period than would direct extrapolation of the periods.

2.11 Migration

A number of new problems enter with the recognition of migration, as Rogers [46, 47] and Lowry [33] have shown. Think of a matrix $\boldsymbol{A}$ whose typical element a_{ij}, $i\neq j$, is the fraction of the population in the jth region at the beginning of each time interval who move to the ith region during the interval. The diagonals a_{ii} are the growth ratios of the several regions due to birth, death, and migration from outside of the whole set of regions.

The matrix $\boldsymbol{A}$ premultiplies the initial vector containing the numbers in the several regions placed one under the other, and the result is the numbers, region by region, at the end of one time interval. As with the age matrix we can continue the multiplication to project into later periods, and the tth power of $\boldsymbol{A}$ contains such terms as $a_{11}^{(t)}$, the fraction of the original number in the first region that will be in the first region after t periods, some having come from elsewhere.

For ordinary projection we need to recognize age and sex within each region. The idea of combining the transitions among regions with those among ages into a larger matrix is found in [37, 46, 16].

2.12 Labor Force

Just as individuals of a given age are classified by region, and may move among regions, so they may be identified with a labor force status and an occupation, with movement among these permitted [6]. The main work on this is by Tabah [54], who developed formulas for the elements of the large matrix, and then applied them to the transition for Mexico from 1960 to 1985. With many occupation, region, and age classes the matrix can become cumbersome, but modern equipment can handle large matrices and computing difficulties need not hold back the development of the subject.

Usually the labor force projection is based on a population projection by age and sex, on which is superimposed a set of participation rates, defined as the fraction of each sex-age total participating in the labor force. If the fractions are represented by a horizontal vector $\boldsymbol{F}$, and this is fixed through time, then the projected total labor force at time t is $\boldsymbol{F}\boldsymbol{M}^t\boldsymbol{P}^{(0)}$. The labor force distributed by age and sex as a vertical vector is given by the same formula, but now the set of participation rates $\boldsymbol{F}$ is incorporated as the diagonal elements of a matrix whose off-diagonal elements are zero.

3. PROJECTION AND FORECASTING

So far we have discussed only projection, defined as the numerical elaboration of a set of assumptions made in order to illuminate data referring to the present or past. The contrast with prediction or forecasting, assertion of what will occur in the concrete future, seems sharp. At the worst the population projection strikes technical difficulties, and most of these are easily superable; time enters only as a variable, the techniques for handling which are readily available, and error is virtually impossible. But forecasting refers to real time, and here we encounter the much deeper difficulties associated with the variability of human behavior; error is practically certain. The difference between elaboration of arbitrary sets of assumptions and prediction of what will really happen in the future is great enough, it would seem, to leave no ground for confusion.

Yet in fact projections and forecasts are not easily distinguished. Projections would hardly be made and published if no one used them as guides to the future. A demographer makes a projection, and his reader uses it as a forecast; does the demographer's intention or the reader's use determine whether projection or forecasting has occurred? Of all the ways in which projections might be made, why are some often used and others not used at all? If projection was merely working out the numerical implications of arbitrary assumptions, then one projection would be as good as another, provided arithmetic errors were avoided. There would be no room for discussing whether we should apply the present age-specific rates or crude rates, for example. In fact no one admits to making an arbitrary choice of assumptions for population projections; each author selects a set corresponding to the relations that he sees as persisting into the future. Since what will persist is uncertain at the time the projection is being made, he is well advised to try more than one set of assumptions and work out future numbers from each. In due course censuses will be taken in what was the future at the time the projection was made. The hope is that the several future numbers will turn out to straddle each subsequent census, but official agencies, unwilling to present their projections as predictions, do not assign any probability that they will straddle.

3.1 Forecasting in Fields Other than Demography

Similar governmental and professional reticence in regard to the future is not shared by other disciplines. Weather forecasters brazenly forecast tomorrow's rain or shine, and users act on the forecast. They are less certain that tomorrow will bring rain than demographers are that next year's population will be larger than this year's in the United States. Even the amount of increase of population is known more precisely than is the amount of rain.

Closer to our problem, economic forecasting is frankly accepted as a main objective of professional and governmental work [25]. Scores of economic forecasts are made each year in respect of the succeeding year or years; a substantial fraction of the reputable members of the economics profession go on record, some of them under the auspices of a government agency that sets an official imprimatur on a scholarly authority.

So little opprobrium attaches to prediction or forecasting in statistical circles that writers profess to be doing it even when they are in fact doing something much simpler; they talk of "predicting" from a regression equation, where what is "predicted" is merely the points on which the regression was based, or other points drawn at random from the same population. If the data come from the past, and the inference or prediction concerns the future, we can never assume that data and inference apply to exactly the same population. It is the component of error arising because the future will be genuinely different from the past that is the characteristically awkward and puzzling feature of prediction. Bell [3] has initiated discussion of the logic of the prediction problem.

3.2 Statement of Assumptions

Notwithstanding the example set by meteorologists, economists and statisticians, demographers making official projections insist they are not predicting. Presumably accurate prediction would require occult powers, which demographers attached to government agencies are not expected to claim. (Anyone who thinks he has such powers can find adequate scope for their exercise in the private sector, for instance in stock market speculation, where he can take financial responsibility for his assertions. Occult powers are either too precious or too chancy for government use.)

In the division of responsibility the official demographer works out the consequences of several sets of assumptions, and leaves it to others to decide which to use. Published projections are often at fault, however, in not making clear exactly what their assumptions are. If the reader is to choose among assumptions, (and not merely to choose among future populations) he must know what the assumptions are, and how those for one projection differ from those for another. Think of how confused the reader of a mathematics article would be if the writer scattered the assumptions underlying a theorem over many pages, some before the theorem and some after, never identifying them as such but burying them in a diffuse text, and stating them in such vague terms that even the reader who was able to identify them could not trace their effect on the theorem. Many published projections do just this.

The point is important enough to be worth exemplifying with a relatively precise description. The following is proposed as an example of style only.

> Two sets of projections for Costa Rica are contained in this report, labelled A and B. The initial age distribution in both was the official estimate by the Bureau of Statistics for January 1, 1969. Following are the assumptions under which they were worked out:
>
> A. 1. Age-specific death rates are those of Costa Rica in 1969, applied without change to all subsequent years, for the two sexes separately.
>
> 2. Birth rates for girls specific for age of mother are those of Costa Rica in 1969, decreasing by 1 percent per year until 1980, then decreasing by 2 percent per year until the year 2,000. Birth rates for boys specific for age of mother are in each year 1.05 times the birth rate for girls.
>
> 3. Net immigration is taken as 5,000 per year, distributed by age and sex as the total immigrants of 1969.
>
> B. 1. Age-specific death rates are those of Costa Rica in 1969, extrapolated for each sex separately by holding constant the first difference between the rates of 1969 and 1959 for each year. Life tables are made from these extrapolated rates as averaged over the five year periods.
>
> 2. Age-specific birth rates are made to conform to a total number of children born by each cohort. The total is assumed to be 5.0 for the cohort of women born in 1940, diminishing by 0.05 for each year thereafter until it reaches 3.5 for the women born in 1970, and thence remaining constant. The total is distributed by age according to the age-specific rates in 1969. The reader will notice discontinuities, especially for those cohorts that have borne most of their children before 1969.
>
> 3. Net immigration is taken as 5,000 per year, distributed by age and sex as the total immigrants of 1969, the same as in projection A.

This imaginary example is given for form of exposition, and not for its substance. In a real piece of work the writer would probably have some way of handling the discontinuities (under B.2) for those women whose families are largely but not entirely complete at the start of the process. Other improvements in the foregoing sets A and B will quickly come to mind—that is the advantage of a formal display.

For such sets of assumptions to be fully understood by the reader without having to search outside, the writer ought to show the data referred to. He ought for instance to show a table of age-specific rates for Costa Rica in 1969 as part of the explanation of (1) in projection A.

To sum up, the official view is that projections are merely the working out of a set of assumptions among which the reader will choose, and his choice constitutes the prediction. This can make good sense if the official text states the assumptions with perfect clarity. They may or may not be scattered through the exposition, but in any case they should be collected together and displayed in clear type. The frequent failure to do this prevents the reader from exercising his judgment in *choosing among the assumptions*; it means that the reader is instead compelled to *choose among the future populations* provided to him. Logically this latter is no different from compelling him to choose a future population from among a set of random numbers.

4. FORECASTING METHODS AND OBJECTIVES

The components method of projection provides forecasts once its assumptions are reinterpreted as representing the real state of affairs. Its several components then receive quite different emphasis. For small areas migration becomes the key variable; for a country as a whole migration is both a smaller fraction of population, and better recorded, than it is for states and counties. Deaths have been the key variable for poor countries in the past, and births for rich countries. In any particular projection one of the three—migration, death and birth—is likely to demand special attention.

This section will review briefly some methods of forecasting, with emphasis on several that have not yet been tried in demography but seem potentially useful.

4.1 Leading and Lagging Series

The view that economic life runs in cycles, and that among the hundreds of available series some reach the turning points earlier and some later, suggests the use of the earlier series to predict the later. If overtime work in manufacturing tends to turn downward from a cyclical high six months before unemployment turns up from its low point, then the overtime series can predict the unemployment series, on the method developed by the National Bureau of Economic Research. Unfortunately little economic theory is available to support the selection of leading and lagging series, and so we cannot be sure of the stability of the sequence [18]. Notwithstanding this criticism, population forecasting would benefit if its variables, and births in particular, could be located in the sequence of indicators. Easterlin [14] has studied births in such a context. The approach is limited to short-term forecasts; the attempt to see years and even decades ahead to which the demographer is pushed by his clients requires other techniques.

4.2 Autoregressive Series

A more abstract study of time series has developed over the last thirty years, whose extensive theory is made conveniently accessible by Box and Jenkins [8]. Time series may be characterized by the correlation between successive terms, between terms separated by one other term, by two other terms, etc. These auto-correlations permit the calculation of multiple regressions by which a term is predicted from the preceding terms; the calculation takes account of all persistent peculiarities of the series.

The most useful part of the theory is concerned with stationary time series, in which the extrapolation from one point in the curve to another depends only on the time interval between them. Unfortunately for application to demography, each decade seems to see historic and more or less unprecedented population changes, especially in births, that are seriously inconsistent with the stationarity assumption. Nonetheless the treatment of sections of the population or birth curve as stationary could at least provide lower bounds to the error of prediction, even if it cannot greatly improve the prediction itself.

4.3 Regression on Other than Population Variables

Insofar as population changes are caused by changes in such variables as income and education, predictions of these latter provide a lever for predicting population. Adelman [1] works out multiple regressions among countries, in which age-specific birth rates at practically all ages turn out to be related positively to income and negatively to education, percent of labor force outside of agriculture, and population density. We can safely forecast increasing proportions outside of agriculture, increasing income, etc., in many parts of the world, and hence ought to be able to conclude that births will be in accord with the higher levels of these other variables. But more study is needed of the conditions under which relations that hold in a cross-section of countries taken at one moment of time hold longitudinally for an individual country.

An econometric model of human fertility, provided by Phillips, Votey, and Maxwell [45], has potential for demographic use. Largely beyond regression analysis are the cultural and institutional settings in which fertility rises and falls; the forecaster cannot neglect these, but neither in the present state of knowledge can he take them precisely into account.

4.4 Another Markov Chain Model

Fish and Thompson [17] make use of the relation of the birth rate to other social variables in a different way, reminiscent of the demographic treatment of age and sex. The lower birth rate in cities as compared with the countryside seems to persist through time irrespective of changing proportions in the city and countryside; can one not project the trend in rural-urban residence and similar variables and then suppose fixed birth rates in each type of area? The Fish-Thompson model does just this. A Markov chain is constructed with fixed transition probabilities for residential and other movement; to the changing composition projected by the Markov chain are applied the birth rates of the initial period.

To take a simplified example, if a_{ru} is the proportion of rural dwellers who move each year to urban parts, and a_{ur} the proportion of urban dwellers for the reverse movement, then the transition matrix is

$$\boldsymbol{A} = \begin{bmatrix} a_{uu} & a_{ru} \\ a_{ur} & a_{rr} \end{bmatrix}$$

where $a_{uu}+a_{ur}=1$ and $a_{ru}+a_{rr}=1$.

If $\boldsymbol{A}$ continues to apply through successive time intervals, and the initial distribution is p_u persons in urban places and p_r in rural, arranged in a vector

$$\boldsymbol{P}^{(0)} = \begin{bmatrix} p_u \\ p_r \end{bmatrix},$$

then the distribution t years later will be

$$\boldsymbol{P}^{(t)} = \boldsymbol{A}^t\boldsymbol{P}^{(0)}.$$

If the overall birth rates in rural and urban parts are b_u and b_r respectively, arranged in a horizontal vector $\boldsymbol{B}$, then the births at time t are

$$\boldsymbol{B}\boldsymbol{A}^t\boldsymbol{P}^{(0)} = [b_u \quad b_r]\begin{bmatrix} a_{uu} & a_{ru} \\ a_{ur} & a_{rr} \end{bmatrix}^t \begin{bmatrix} p_u \\ p_r \end{bmatrix}. \tag{4.1}$$

In effect (4.1) extends to other variables the specificity by age and sex of the process of projection described above, for example in (2.6).

4.5 Survey of Childbearing Intentions

A direct way of getting at the future is to ask people what they intend to do. Firms are surveyed on their investment intentions and their employment plans, and these surveys play an important part in economic forecasting. Ever since World War II individuals have been asked whether they intend to buy or build a house, or to buy automobiles and smaller consumer goods, and these also have provided indications of future economic activity [9]. Why not ask couples whether they intend to have a child? This is now done regularly, and provides useful information for a short period ahead, say about five years. Beyond that time most of the babies born will be to women not yet married at the time of interview, who can hardly have formed any intention firm enough to be communicated to a survey enumerator [52, 50].

4.6 The Demographic Transition as a Repeated Historical Pattern

The course of population growth in advanced countries has included a fall in the death rate, followed at a greater or lesser interval by a fall in the birth rate. Before the first event and after the second the population grows slowly; between the two events it grows very rapidly. This relation should be usable for forecasting. Indeed it may be said that many of the achievements of population prediction, especially Cannan's original use of the method of projection by components, occurred when western countries were in the declining birth rate phase of the transition. Application of the theory is aided by a finding of Dudley Kirk that the rapidity of the fall in the birth rate increases with the lateness of the onset of the fall; Puerto Rico in the 1950's showed a fall five times as steep as did Holland or England in the late 19th Century [24].

Although the demographic transition is a thoroughly documented description of what has happened in the past, and no one can forecast populations of less developed countries without keeping it in mind, yet the variety of circumstances in its past history leaves us unsure as to just when and how it will happen in individual countries. One nagging uncertainty is the relation of the fall of the birth rate to economic development. That relation has not been uniform—during the nineteenth century England's income rose faster than France's, but its birth rate fell more slowly. Yet the countries whose birth rate has started to fall in recent decades, such as Taiwan and Singapore, are those of relatively high income. We cannot either suppose that income is necessary for the fall of the birth rate nor that the birth rate falls according to a dynamics of population transition that is independent of income. The problem goes beyond population forecasting into the mechanics and strategy of economic development.

4.7 Forecasting an Ecological Ceiling

Any theory of ecological limits to population growth is an aid to forecasting once high densities are reached. The exponential curve is often called Malthusian, but Malthus thought of population as increasing geometrically only in circumstances where nothing stood in the way, and most of the time limitation of resources does stand in the way. Hence the Malthusian fixed resources suggest not geometric increase but a logistic curve in which population rises to a ceiling a:

$$P_t = \frac{a}{1 + e^{-r(t-t_0)}},$$

where r is the rate at which the population increases when it is very small, and t_0 is the time when it reaches the half-way mark on the curve, i.e., a population $a/2$. Such a curve was used by Pearl and Reed [44], but after some initial success fell into discredit because it overshot the 1940 census. The limitations of resources seen by present-day ecologists could bring the logistic or similar curves back into fashion as a forecasting device.

4.8 The User's Loss Function

Population projections are ordinarily published by demographers for general use by planners, enterprises, and others. But each use requires its own degree of accuracy, and each stands to lose in a different way as a result of inaccuracy. A definition and application of the loss function [40] clarifies the division of responsibility between the producer and the user of population projections.

If the population at some future date is taken by the user as $\hat{x}$, and the true population at that date turns out subsequently to be x, let $L(\hat{x}, x)$ be the loss to the user for this particular combination. Considered as a function of x the function $L(\hat{x}, x)$ will be minimum when the true future x is the same as the anticipated $\hat{x}$, but it need not be symmetrical about $\hat{x}$. If $\hat{x}$ is the expected future population of a town for which a water reservoir is to be designed, then an underestimate (i.e., x turning out larger than $\hat{x}$) may mean the very large cost of building a new additional reservoir, while an equal overestimate may mean only some underutilization—the reservoir was built too large, but this did not add much to the original cost.

To apply his loss-function the user needs a probability distribution of the future population x. Suppose that this is provided by the demographer as the function $P(x)$, which says that the probability that the population will turn out to be between x and $x+dx$ is $P(x)dx$. For any x the expected loss for the particular user will be $L(\hat{x}, x) \cdot P(x)dx$, and the total expected loss will be $R(\hat{x}) = \int_{-\infty}^{+\infty} L(\hat{x}, x)P(x)dx$, a function of $\hat{x}$ from which x, the true future population, has been integrated out. The user wants the value of $\hat{x}$ for which this quantity $R(\hat{x})$ is smallest, and he finds it by equating the derivative to zero,

$$\frac{dR(\hat{x})}{d\hat{x}} = 0 \tag{4.2}$$

and solving for $\hat{x}$. The simple equation (4.2) would provide the right future population $\hat{x}$ for the particular use.

To permit the calculation of $\hat{x}$ the demographer must produce not one or several estimates, but a probability distribution $P(x)$ for each future date. A mean and variance for each future date would suffice, since $P(x)$ could reasonably be taken as normally distributed. In practice the several projections produced officially are now applied in somewhat this way, at least in that a user especially afraid of underestimating would take the largest of the projections officially provided, and one more afraid of overestimating would take the lowest. The explicit probability distribution would encourage the user to give more attention to his own loss function.

4.9 The Objectives of Future Population Estimates

Several kinds of future have necessarily entered our discussion, but these are far from inclusive. Projections, predictions and minimum loss population for the purpose of a particular project, are all members of a larger set of modes of reference to the future. To locate the previous discussion in a wider context, I submit that among other objectives, future population may be calculated as

1. A *projection*, working out the numerical consequences of a set of assumptions, dealt with at length above. Several projections may be made at the same time by the same author, it being left to the user to select the set of assumptions that appear to him most realistic.
2. A *forecast*, preferably given as a probability distribution. One might assert that the population of the United States in 1980 will be centered on the range of 225 to 235 million with probability 2 to 1, as though a probability is assigned to each possible future number in a normal distribution with mean 230 and standard deviation 5 million.
3. An industrial or government planner's *estimate for the purpose of a particular project*. As we have seen, the planner's loss function superimposed on the forecaster's probability distribution can provide a minimum loss estimate.
4. A *counter-prediction*—the assertion that something cannot happen. Mexico cannot continue to increase at the rate of 3.5 percent per year, for this would cause it to double every 20 years, and thus multiply by 32 over the next century. The resulting 1.6 billion Mexicans in the present space, if not a physical impossibility, is certainly an economic one.
5. *Normatively*, a target. Pakistan aims to reduce its birth rate to 30 per thousand by 1975. Some feel that the United States population ought not to be allowed to exceed 300,000,000, or that zero population growth ought to be reached by the turn of the century. Once the target is given the demographer is invited to calculate the smoothest, least painful, or minimum cost trajectory for attaining it; this is the converse of the projection problem of 1.
6. *Conflicting estimates* from different sources. Labor force employed is the prime example. We can estimate the supply of labor from knowledge of the population of ages 15–64, applying certain participation rates. On the other hand the demand for labor depends on industrial and governmental activity. *Ex ante* the two will not in general be the same; *ex post* they have to be the same. The separate estimates of supply and demand invite one to foresee how they will be reconciled in the future actuality [20].

Especially in application to 2. above, a sensitive forecaster will be aware of variations that he might indifferently have introduced at the several stages of his calculation, and so will be able to come up with a range on whose straddling of the subsequent census he would be willing to place a 19:1 bet, and which on a normal distribution is equal to four standard deviations. Such estimates of error may be called *ex ante*, in a somewhat different usage from that of 6. in the preceding list. On the other hand the comparison between the forecast and the subsequent census realization ought to provide an *ex post* estimate of error. When a number of *ex post* errors have been accumulated for a given method of prediction they can help the forecaster verify the level of his *ex ante* errors, or else give him a correction factor for them. *Ex post* error for a specimen instance is calculated in Section 6.

5. OFFICIAL ESTIMATES OF FUTURE POPULATION

We proceed to cite some of the projections that have been made for the future population of the world and of the United States.

5.1 The Series of Estimates for World Population in 1980

The global population projections of the United Nations [55, 56, 58] dating back to 1951, have been compiled with much care. ([34, 15, 35] give recent methods and results.) The series starts with relatively simple estimates that include alternative grand totals for 1980 of 2,976 and 3,636 millions (Table 1).

1. ESTIMATES OF 1980 WORLD POPULATION[a]

Made in	Low	Medium	High
1951	2,976		3,636
1954	3,295		3,990
1957	3,850	4,220	4,280
1963	4,147	4,330	4,550
1968		4,457	

[a] Millions of persons, as made by the United Nations at several points of time from 1951 to 1968.

With the release of the 1950 censuses and a higher estimate for the 1953 population of China the United Nations demographers saw that their numbers for the jumping-off dates of the 1951 projections were low. More elaborate projections were made in 1954 showing separately 25 regions into which the map of the planet was divided. Now the lowest figure for 1980 was 3,295 million and the highest was 3,990 million. By 1957 a further set was made, in which the lowest had become 3,850 million and the highest 4,280 million, still for the year 1980.

In 1966 the United Nations published its assessment of future population made in 1963 [56], and again the curves were raised; the low figure for 1980 had gone up to 4,147 million, and the high figure to 4,550 million. A further considerable rise was shown by 1968—the medium variant for 1980 had risen from 4,330 million as estimated in 1963 to 4,457 million as estimated in 1968. These estimates for 1980 are summarized in Table 1. It is reasonable that the range from low to high should diminish (from 660 million in 1951 to 403 million in 1963 in this case); as the date of 1980 approaches the range of uncertainty in its estimate ought to narrow.

What is more striking is that the Medium Variant 1980 projection had gone from about 3.3 billion to 4.5 billion over the 17 years. The increase was 36 percent or about 1.8 percent per year, a growth of the same order as the population itself showed over the period of the estimates! This arose because the national estimates on which the United Nations had to rely were shown to be gross underestimates by censuses taken in 1960 as well as 1950, and death rates fell much faster during the 1950s and 1960s than could have been anticipated.

Such comparison of projections made at different times for the same future date may give some feeling for the magnitude of error to which estimates of future population are subject. Less than a decade suffices for the ranges to become non-overlapping: the high estimate of 1954 is lower than the low estimate of 1963 (Table 1). The alterations were greater for the less than for the more developed regions. Between 1963 and 1968, the less developed went from a 1980 prospective population of 3,136 million to 3,247, an increase of 3.5 percent, while the more developed regions went from 1,194 million to 1,210 million, an increase of only 1.3 percent, all on the medium variant.

5.2 The Year 2000

Let us turn now to the estimates for the year 2000 as assessed in 1963, the most complete set so far. The low figure for the year 2000 is 5,449 million, and the high is 6,994 million. This wide range—roughly 5.5 billion to 7.0 billion—is still about the closest that anyone can yet come to the total for the year 2000. This writer would give nineteen to one odds that the true 2000 figure will fall in the range 5.5 to 7.0 billion, but not for any narrower range. The medium variant, 6,130 million, is probably on the low side, if only to judge from the fact that the 1968 medium variant was higher. All of the figures mentioned assume some drop in fertility: projections with constant fertility are high above these.

5.3 The World Rate of Increase and Its Imminent Decline

The medium estimates made in 1963 show an increase in the annual absolute increase up to the year 2000. The rise from 1995 to 2000 is 482 million, more than half as large again as the latest past increase of 311 million in 1965–70. But at least the increase of the increase is falling. The percent increase over five-year intervals also falls after reaching a peak of 9.80 percent in 1970–5. Insofar as it is the percent increase whose acceleration through the last few centuries constitutes the population explosion, we can obtain some comfort from its peaking in the 1970's. At least in this respect we are at an historic turning point, even though absolute increase will keep climbing.

The reason for the peak and then decline in the rate of natural increase is contained in its relation to death rates. Natural increase is in the long run unaffected by changes in mortality beyond the ages of reproduction, and mortality has tended historically to fall first and most rapidly at the ages of reproduction and below. Once the probability of surviving to about age 40 has risen to about 0.9 further decline of mortality can have little further effect on long term population increase. At this point even a small fall in fertility will cause the rate of increase to turn downwards, and fertility in the world as a whole is falling, even if slowly. Hence the imminent peaking of the rate of increase for the world as a whole.

5.4 Contrast Between Regions

Continuing with the 1963 assessment, we find a large difference between the more and the less developed regions:

	1970	2000
	Millions	
More developed regions	1,082	1,441
Less developed regions	2,510	4,688

During the 30 years the less developed go from 2.3 times the more developed to 3.3 times.

This corresponds to striking variation among continental areas. At the lowest, Europe increases from 454 million in 1970 to 527 million in 2000, or only one sixth. This is followed by East Asia (mostly Japan and Mainland China) which increases by about two-fifths, the USSR which increases by somewhat more than this, and Northern America which increases a little more than half. In contrast to these South Asia nearly doubles, and Africa and Latin America more than double during the 30 years. The population of Latin America first climbed past that of Northern America in the 1950's, and by the year 2000 it stands at 638 million against 354 million for Northern America.

John Durand [13] and Jean Bourgeois-Pichat [5] put some of these numbers in a longer perspective. They estimate that the population of the countries that were developed in 1960 (not quite equivalent to the countries of European origin) were 20.4 percent of the world total in 1750, rose to a peak of 30.9 percent of the world total in 1900, then declined to 28.6 percent in 1960 and by 2000 decline further to 21.3 percent or nearly back to their percentage in 1750.

We cannot deal with each individual country here, but

2. UNITED NATIONS 1963 MEDIUM VARIANT ESTIMATE OF WORLD POPULATION AT 5-YEAR INTERVALS, 1960–2000[a]

Year	Total	Absolute increase	Percent increase
1960	2,998		
		283	9.44
1965	3,281		
		311	9.48
1970	3,592		
		352	9.80
1975	3,944		
		386	9.79
1980	4,330		
		416	9.61
1985	4,746		
		442	9.31
1990	5,188		
		460	8.87
1995	5,648		
		482	8.53
2000	6,130		

[a] Millions of persons.

show the eight countries of largest population in 1970 (Table 3). Combined these constitute over 60 percent of the world total. Four of them, Japan, the United States, the USSR, and surprisingly Mainland China, show low rates of increase, from 10.0 to 15.9 percent for the decade 1970–80, again according to the UN medium projection of 1963. The other four countries range from India's 25.6 percent up to Pakistan's 36.6 percent.

3. EIGHT LARGEST COUNTRIES IN ORDER OF 1970 POPULATION, SHOWING PROJECTION TO 1980 AND PROJECTED RATE OF INCREASE[a]

Country	1970	1980	Percent increase
Mainland China	742	843	13.6
India	543	682	25.6
USSR	246	278	13.0
United States	208[b]	241	15.9
Pakistan	134	183	36.6
Indonesia	118	153	29.7
Japan	101	111	10.0
Brazil	94	124	31.9
Total 8 countries	2186	2615	19.6
World	3592	4330	20.5

[a] Millions of persons, as assessed by the United Nations in 1963, Medium Variant.
[b] Calculation antedates the 1970 U.S. Census count.

5.5 Urbanization

The extraordinary growth of cities is brought out in a further elaboration of the United Nations (1963) projections. Between 1920 and 1960 world population increased by 61 percent, while the population of localities of 20,000 and more inhabitants increased by 185 percent. This growth of cities was higher in percentage terms for the less developed (326 percent) than for the more developed (116 percent), and even in absolute terms the more developed added 209 million, the less developed 284 million [57]. In the 19th century city populations were associated with industrialization, but this association does not always hold in the 20th century.

The larger growth of cities in the underdeveloped areas appears even more strikingly in United Nations projections for the period 1960–2000. Now the places of 20,000 and over in developed areas go from 389 million to 784 million, and in less developed areas from 371 to 1,553 million. By the year 2000 there will be about twice as many large-city inhabitants in underdeveloped as in developed countries. Considering present rates of capital accumulation in the two groups of countries this implies a striking change in the nature of urbanization, and especially the way in which urban residents of poor countries obtain their livelihood.

5.6 The United States in the Year 2000

If we are uncertain of future United States population it is not because forecasts have been lacking. This discussion will be confined to the careful projections made by the Bureau of the Census, in particular the set released in August 1970, and carried in the 1970 *Statistical Abstract* [61]. Four series are shown, of which the highest, labelled B, stands at 321 million for the year 2000, and the lowest labelled E at 266 million. The four replace estimates dated December 1967, which range from Series A, 361 million, down to Series D, 283 million (Table 4). The 1970 publication differs from that of 1967 mainly in that the A Series was dropped, and a new Series E was introduced.

4. PROJECTION OF POPULATION OF THE UNITED STATES TO THE YEAR 2000[a]

Projection made in	A	B	C	D	E
1967	361	336	308	283	
1970		321	301	281	266

[a] Millions of persons.
Source: U.S. Bureau of the Census, Series P-25, No. 381, December 1967, Series P-25, No. 448, August 1970.

The several series are distinguished chiefly by their fertility assumptions, most conveniently expressed in terms of completed family size for women living to the end of childbearing: Series B supposes 3.10 children and Series E 2.11 children. The difference of about one child per couple, for those couples starting their childbearing now, is what makes the difference between 266 and 321 million in the year 2000. For those couples that are already embarked on their childbearing careers the projections take account of the children born to date, and suppose a gradual approach to the ultimate averages shown in Table 5.

5. AVERAGE TOTAL COHORT FERTILITY AND TIMING ASSUMED IN U.S. PROJECTIONS A TO E

Series	Ultimate total children born per woman	Ultimate median age of mother
A	3.35	25.3
B	3.10	25.8
C	2.775	26.4
D	2.45	27.2
E	2.11	25.8

The determination of the ultimate numbers of children per woman in Table 5 is based on past history and on surveys of birth expectations. The women who did most of their bearing in the 1950's had somewhat more than 3.35 children on the average, and at the other extreme 2.11 is about the average that will ultimately make a stationary population. A rough way of seeing this latter point in terms of the female dominant model is to take the fraction of births that are girls, about 0.489, so that the 2.11 children is equivalent to about $2.11 \times 0.489 = 1.032$ girls per woman. The required allowance for mortality to translate this last number into replacement is the ratio of the Net to the Gross Reproduction rate (NRR to GRR), which in recent years has been about 0.96. Multiplying 1.032 by 0.96 gives 0.99 girls arriving at reproductive age to replace one now, i.e., a practically stationary population.

If the E estimate assumes bare replacement or less, why does the E population continue to rise, not only until the year 2000 but for twenty or more years thereafter? The answer is in the age distribution of the United States, which has been determined especially by the large number of post-war births, and so contains a higher proportion of women of childbearing age than the ultimate stationary population will. While these are having their families, even though each averages only two children living to maturity, the birth rate will be higher than the ultimate stationary rate.

To ask which of the five series is preferred if one must choose a single number is to move from the relatively safe ground of projection to the hazardous one of prediction. My own preference at this time is estimate D: if I had to give one number I would say 280,000,000 for the United States in the year 2000. But I would be unwilling to lay 19 to 1 odds on any range narrower than 240 to 320 million; my subjective probability distribution for the year 2000 is normal with mean 280 million and standard deviation 20 million. Insofar as this impressionistic statement has an empirical basis, it is observation of the errors of predictions by competent demographers in the past.

Even more hazardous is an estimate of where, within the United States, population will be concentrated. California was the goal of much internal migration during the postwar period. The Bureau of the Census projects for it 28 million persons in 1985, based on migration patterns of 1955–60 [61, p. 13]. This increase of about 40 percent is far more than shown by most other states; the West North Central and East South Central areas are nearly stationary in the projection to 1985. Should we judge California's prospects by the 1955–60 migration pattern, when aircraft and space industries were expanding rapidly, or by the late 1960s and early 1970s when these industries are being in part dismantled, and California's unemployment rates are higher than the average of the country? The smaller the area the more closely is its future population linked to economic prospects.

At several points I have made assertions in terms of a subjective probability distribution of the future, expressed as a range on whose straddling the population of the year 2000 I would give 19:1 odds. The demographer should be encouraged to such subjective probabilities by the thought that society bets on future population whenever it builds a school, a factory, or a road. Real wagers, running to billions of dollars, are implicit in each year's capital investment. Someone ought to be willing to make imaginary wagers if so doing will more precisely describe the distribution of future population and so improve investment performance even by a minute fraction.

When a forecast is in the form of a distribution the user sees immediately how much he can trust its mean. He is entitled to the warning constituted by a wide distribution that the forecaster is unsure of his statement of the mean; the least that can be said for the subjective statement of variance is that it is a compact way of expressing the forecaster's uncertainty. In fact it is much more than that; after the time for which the forecast is made has come around it permits observation of where the realization fell in relation to the forecast distribution Evaluation after the event is no problem for forecasts that show variance as well as mean.

It is a problem for point estimates, and in Section 6 we consider how point-forecasts can be evaluated after the event. The problem is to devise an objective procedure for what has above been called *ex post* error.

6. EX POST EVALUATION

Even after the event a statement of the accuracy of prediction is not easy. A bench mark of some kind is needed analogous to the "persistence" (forecasting that tomorrow's weather will be the same as today's) that meteorologists use to see how well their elaborate methods are serving and that Mosteller [39, p. 297] has applied to election polls.

For when a forecast is a single number and not a probability distribution its error can only be assessed by comparison with some standard. The Scripps medium projection made in 1935 for the 1970 population of the conterminous United States was 155 million [63, p. 470]. Let us treat this as a forecast for purposes of this illustration. The performance as counted in the 1970 census was 203 million for the same area. The question whether the forecast was good or bad cannot be answered in absolute terms; it is meaningless to ask if 48 million is a large error in general. We can note that the estimate was 24 percent short, and can only say that if one could predict the national income of the year 2100 within 24 percent he would be doing well; to be in error 24 percent in predicting next year's Federal tax collections would be decidedly poor.

Though the quality of a prediction must necessarily be assessed in relation to some other prediction produced by a standard or naive method that can be taken as a base or bench mark, it is not obvious how that base should be selected. One way would correspond to persistence in weather forecasting: that the population will not change from the 1935 level. The 1935 level was 127 million, so the true increase over the whole period 1935–70 was $203 - 127$

$=76$ million. This was the error to which the use of the 1935 figure for 1970 would have been subject. Of this the 1935 medium projection embraced $155-127=28$ million. Thus the 1935 projection was an improvement on the persistence error of 76 million; it avoided 28 million of the error to which one would have been subject in using the 1935 population. Its score may be given as $28/76=0.37$.

The general formula for the quality of a prediction as measured *ex post* is

$$\text{Quality of Prediction} = \frac{\text{Prediction—Bench mark}}{\text{Realization—Bench mark}}.$$

The simplest is to use a bench mark of zero, giving for 1970

$$\text{Quality} = \frac{155-0}{203-0} = 0.76.$$

A second method is to use the 1935 observed population as a bench mark, giving, again for 1970

$$\text{Quality} = \frac{155-127}{203-127} = \frac{28}{76} = 0.37.$$

We now proceed to a yet tighter measure of quality. The forecast will no longer be credited with the fraction of the increase from the jumping-off point that it includes, but only with that part of it not covered by a projection with fixed 1935 rates of birth and death:

$$\text{Quality} = \frac{\text{Prediction—Projection at fixed rates}}{\text{Realization—Projection at fixed rates}} = \frac{155-148}{203-148} = \frac{7}{55} = 0.13.$$

This illustration is typical of the fact that carefully made projections by professional demographers come closer to the target than fixed rate projections, but the improvement is not striking.

7. CONCLUSION

That the future is important to everyone, that many good minds have worked to produce a methodology for foretelling it, that many hundreds of professionals are working to produce numbers saying what future population might be, all of these do not add up to any great degree of certainty even for short-term projections, let alone for long-term [11, 12, 19].

The preceding pages have included some suggestions for improving performance. They draw attention to a variety of methods for projection and prediction, not all of them yet applied to future population. Practitioners might well select more widely. As a matter of exposition, some of current publication is lacking in clarity on just what its assumptions are. The ideal is a clear enough statement that the reader could reproduce the numbers of the projection on his own desk calculator or computer if he wanted to badly enough. Such an ideal may be unattainable but it could be usefully approached more closely than in most current practice. Clarity of exposition would encourage professional discussion, and out of this better methods should emerge. In deliberations on future population one should study not only the latest estimates of official sources, but also their earlier estimates, to gain some impression of forecasting errors. The aim is (necessarily subjective) probability distributions of population at each future date.

This article has gone from a firm starting point in projection, through the unmarked wilderness of prediction, in which one is left to his virtually unguided intuition, and on to evaluation when the prediction is checked by a subsequent census. We know how to make projections with great precision once the assumptions are chosen, but projections as such exercise considerably less fascination for the user than they do for the scholar who makes them. Population forecasts are demanded as strongly as forecasts of the weather or national income.

Given that the responsibility of official agencies prevents their making forecasts, and that the user typically claims incompetence, there seems to be room for a middleman, who will take the projections and apply his judgment to convert them into forecasts. If he provides the forecast in the form of a probability distribution, as is strongly recommended, and the user knows his own business well enough to set up a loss function, then a point forecast is obtained as that future population that minimizes expected loss on the particular project.

The weakness of population forecasts is due to our ignorance of the mechanisms by which populations grow and decline. We know much about birth rates and their differentials among statistically recognizable population subgroups, as well as about changes over time as shown in past records, but this great volume of statistical information has contributed disappointingly little to the discernment of a comprehensive causal system underlying the differentials and changes. My own generalization from postwar experience is that advanced countries will have an endless series of ups and downs in their births, about an average near replacement level. But even if this is correct it would be helpful for prediction only if we had some indication of the wave-length and the amplitude of the cycles.

The projection with fixed rates of birth and death, with which the article starts, may turn out to be a good prediction in circumstances where births move up and down without any net trend. It also serves as a benchmark for evaluating predictions after the event, with which the article ends. This *ex post* error informs us not only of the quality of forecasts, but of bias in the *ex ante* estimates of their error.

The *ex ante* error, calculated on the same data as the forecast itself, may be in the form of a variance, so that it along with the forecast provides a normal (or some other) probability distribution of the predicted population. The *ex post* error contains a component that cannot be part of the *ex ante*· that arising because the future is different from the past. It is the existence of this component that constitutes the essential risk of forecasting.

If the past and the future were one homogeneous series, then calculation of future population would be subject to the laws of sampling, and forecasting would be a routine application of statistical inference. If the future is wholly different from the past, no amount of data and experience can assist prediction. The population forecaster can only suppose that reality falls between these extremes. He is invited to devise appropriate models and apply available statistical methods, but nothing can prevent recurrent challenges to his knowledge in the form of departures of the population realized from that forecast.

[*Received January 1972.*]

REFERENCES

[1] Adelman, Irma, "An Econometric Analysis of Population Growth," *American Economic Review*, 53 (1963), 314–39.

[2] Akers, D. S., "Cohort Fertility versus Parity Progression as Methods of Projecting Births," *Demography*, 2 (1965), 414–28.

[3] Bell, Daniel, "The Year 2000—The Trajectory of an Idea," *Daedalus*, 96 (1967), 639–51.

[4] Bernardelli, H., "Population Waves," *Journal of Burma Research Society*, 31, Part I (1941), 1–18.

[5] Bourgeois-Pichat, J., "Population Growth and Development," *International Conciliation*, No. 556 (1966), New York: Carnegie Endowment for International Peace.

[6] Bowen, W. G. and Finegan, T. A., *The Economics of Labor Force Participation*, Princeton, N. J.: Princeton University Press, 1969.

[7] Bowley, A. L., "Births and Population of Great Britain," *Journal of the Royal Economic Society*, 34 (1924), 188–92.

[8] Box, C. E. and Jenkins, G. M., *Forecasting and Control*, San Francisco: Holden-Day, 1970.

[9] Burch, S. W. and Stekler, H. O., "The Forecasting Accuracy of Consumer Attitude Data," *Journal of the American Statistical Association*, 44 (December 1969), 1225–33.

[10] Cannan, E., "The Probability of a Cessation of the Growth of Population in England and Wales During the Next Century," *The Economic Journal*, 5 (1895), 505–15.

[11] Davis, J. S., "Implications of Prospective United States Population Growth in the 1960s," *Milbank Memorial Fund Quarterly*, 39 (1961), 329–49.

[12] Dorn, H. F., "Pitfalls in Population Forecasts and Projections," *Journal of the American Statistical Association*, 45 (September 1950), 311–34.

[13] Durand, J. D., "The Modern Expansion of World Population," *Proceedings of the American Philosophical Society*, III, *Population Problems*, Philadelphia: American Philosophical Society, 1967, 136–59.

[14] Easterlin, R. A., "The Baby Boom in Perspective," *American Economic Review*, 51 (1961), 869–911.

[15] El-Badry, M. A., "Latin American Population Prospects in the Next Fifteen Years: A Brief Analysis," Conference of the International Union for the Scientific Study of Population, Mexico City, 1970.

[16] Feeney, G. M., "Stable Age by Region Distributions," *Demography*, 7 (1970), 341–68.

[17] Fish, Mary and A. A. Thompson, "The Determinants of Fertility: A Theoretical Forecasting Model," *Behavioral Science*, 15 (1970).

[18] Gordon, R. A. "Alternative Approaches to Forecasting: The Recent Work of the National Bureau," *Review of Economics and Statistics*, 44 (1962), 284–91.

[19] Hajnal, J., "The Prospects for Population Forecasts," *Journal of the American Statistical Association*, 50 (June 1955), 309–22.

[20] Johnston, D. F. "The Integration of Supply and Demand Projections of Labour Force," *Proceedings of the 37th Session, Bulletin of the International Statistical Institute*, 43, Book 1 (1969), 359–72.

[21] Keyfitz, N., *Introduction to the Mathematics of Population*, Reading, Mass.: Addison-Wesley Publishing Co., 1968.

[22] ———, "Finding Probabilities from Observed Rates, or How to Make a Life Table," *The American Statistician*, 25 (1970).

[23] ———, and Flieger, Wilhelm, *Population: Facts and Methods of Demography*, San Francisco: W. H. Freeman, 1971.

[24] Kirk, D., "The Accelerating Decline of Natality in the Demographic Transition," *Population Index*, 36 (1970), 271–3.

[25] Klein, L. R., "Whither Econometrics?" *Journal of the American Statistical Association*, 66 (June 1971), 415–21.

[26] Leslie, P. H., "On the Use of Matrices in Certain Population Mathematics," *Biometrika*, 33 (1945), 183–212.

[27] ———, "Some Further Notes on the Use of Matrices in Population Mathematics," *Biometrika*, 35 (1948), 213–45.

[28] ———, "On the Distribution in Time of the Births in Successive Generations," *Journal of the Royal Statistical Society*, 61, Part I (1948), 44–53.

[29] Lewis, E. G., "On the Generation and Growth of a Population," *Sankhyā*, 6 (1942), 93–6.

[30] Lopez, A., *Problems in Stable Population Theory*, Princeton, N.J.: Office of Population Research, 1961.

[31] Lotka, A. J., "Relation Between Birth Rates and Death Rates," *Science*, N.S., 26 (1907), 21–2.

[32] ———, *Théorie Analytique des Associations Biologiques*, Part II, Analyse Démographique avec Application Particulière à l'Espèce Humaine (Actualités Scientifiques et Industrielles, No. 780), Paris: Hermann & Cie, 1939.

[33] Lowry, I. S., *Migration and Metropolitan Growth*, San Francisco: Chandler, 1966.

[34] Macura, M., "Demographic Prospects for the Next Thirty Years," Unpublished manuscript, 1968.

[35] ———, and El-Badry, M. A., "Diversity or Uniformity of Demographic Problems? A Medium-Range Outlook," in *World Views of Population Problems*, Budapest: Hungarian Academy of Sciences, 1968.

[36] Madrigal, R. E., "Proyecciones de Mortalidad Para Mexico, 1960–1980," Paper submitted to meeting of International Union for the Scientific Study of Population, Mexico City, Aug. 16–21, 1970.

[37] Matras, J., "Social Mobility and Social Structure: Some Insights from the Linear Model," Paper prepared for the Sixth World Congress of Sociology, Evian, France, 1965.

[38] McFarland, D. D., "On the Theory of Stable Populations: A New and Elementary Proof of the Theorems Under Weaker Assumptions," *Demography*, 3 (1969) 301–22.

[39] Mosteller, F., *et al.*, *The Pre-election Polls of 1948: Report to the Committee on Analysis of Pre-election Polls and Forecasts*, New York: Social Science Research Council, 1949.

[40] Muhsam, H. V., "The Utilization of Alternative Population Forecasts in Planning," *Bulletin of the Research Council of Israel*, 5 (1956), 133–46.

[41] Murphy, E. M., "A Generalization of Stable Population Techniques," Unpublished Ph.D. dissertation, Department of Sociology, University of Chicago, 1965.

[42] Oechsli, F., "The Parity and Nuptiality Problem in Demography," Unpublished manuscript, 1970.

[43] Parlett, B., "Ergodic Properties of Populations, I: The One Sex Model," *Theoretical Population Biology*, 1 (1970).

[44] Pearl, R. and Reed, L. J., "On the Rate of Growth of the Population of the United States Since 1790 and Its Mathematical Representation," *Proceedings of the National Academy of Science*, 6 (1920), 275–88.

[45] Phillips L., Votey, H. L. and Maxwell, D. E., "A Synthesis of the Economic and Demographic Models of Fertility: An

Econometric Test," *Review of Economics and Statistics*, 3 (1969), 298–308.
[46] Rogers, A., *Matrix Analysis of Interregional Population Growth and Distribution*, Berkeley: University of California Press, 1968.
[47] ———, *Matrix Methods in Urban and Regional Analysis*, San Francisco: Holden-Day, 1971.
[48] Ryder, N. B., "The Process of Demographic Translation," *Demography*, 1 (1964), 74–82.
[49] ———, "Notes on American Fertility," in N. Keyfitz (ed.), *Statistical Problems of Population Research*, Honolulu: University of Hawaii Press, 1972.
[50] ———, and Westoff, C. F., "The Trend of Expected Parity in the United States: 1955, 1960, 1965," *Population Index*, 33 (1967), 153–68.
[51] Sharpe, F. R. and Lotka, A. J., "A Problem in Age-Distribution," *Philosophical Magazine*, Series 6, 21 (1911), 435–8.
[52] Siegel, J. S. and Akers, D. S., "Some Aspects of the Use of Birth Expectations Data from Sample Surveys for Population Projections," *Demography*, 6 (1969), 101–15.
[53] Sykes, Z. M., "Some Stochastic Versions of the Matrix Model for Population Dynamics," *Journal of the American Statistical Association*, 44 (March 1969), 111–30.
[54] Tabah, L., "Représentations Matricielles de Perspectives de Population Active," *Population*, 23e année (1968), 437–76.
[55] United Nations, *The Future Growth of World Population*, Population Studies No. 28, ST/SOA/Series A/28, New York, 1958.
[56] ———, *World Population Prospects as Assessed in 1963*, Population Studies, No. 41, ST/SOA/Series A/41, New York, 1966.
[57] ———, *Growth of the World's Urban and Rural Population, 1920–2000*, Population Studies, No. 44, ST/SOA/Series A/44, New York, 1969.
[58] ———, *World Population Prospects, 1965–2000, as Assessed in 1968*, Working Paper, No. 37, ESA/P/WP.37, New York 1970.
[59] United States Bureau of the Census, *Current Population Reports*, Series P-25, No. 381, "Projections of the Population of the United States by Age, Sex, and Color to 1990, With Extensions of Population by Age and Sex to 2015," Washington, D.C.: U.S. Government Printing Office, 1967.
[60] ———, *Current Population Reports*, Series P-25, No. 448, "Projections of the Population of the United States, by Age and Sex: 1970 to 2020," Washington, D.C.: U.S. Government Printing Office, 1970.
[61] ———, *The Statistical Abstract of the United States*, 91st edition, Washington, D.C.: U.S. Government Printing Office, 1970.
[62] Whelpton, P. K., "Population in the United States, 1925–1975," *American Journal of Sociology*, 34 (1928), 253–70.
[63] ———, "An Empirical Method of Calculating Future Population," *Journal of the American Statistical Association*, 31 (1936), 457–73.
[64] ———, *Cohort Fertility: Native White Women in the United States*, Princeton, N.J.: Princeton University Press, 1954.

Technical Reports (continued from page 323)

LePage, Raoul D., "Note Relating Singularity of Second Order Processes to Reproducing Kernels," July 1971, *RM-283*.

University of Michigan, Department of Statistics, 1447 Mason Hall, Ann Arbor, Mich. 48104.

Dwyer, Paul S., "Moment Functions of Sample Moment Functions," July 1971, *Report No. 1*.

Raeside, D.E. and Owen, R.J., "A Class of Loss Functions of Catenary Form," July 1971, *Report No. 2*.

Starr, Norman, "How to Win a War if You Must," July 1971, *Report No. 3*.

Kurotschka, V.G. and Dwyer, P.S., "Optimum Design of Three Way Layouts Without Interactions," July 1971, *Report No. 4*.

Bell, von C.B. and Kurotschka, V.G., "Einige Prinzipien zur Behandlung nichtparametrischer Hypothesen," July 1971, *Report No. 5*.

Rothman, E.D. and Woodroofe, Michael, "A Cramer Von-Mises Type Statistic for Testing Symmetry," July 1971, *Report No. 6*.

Starr, Norman and Woodroofe, Michael, "Further Remarks on Sequential Estimation: The Exponential Case," July 1971, *Report No. 7*.

University of Minnesota, Department of Statistics, Minneapolis, Minn. 55455.

Das Gupta, Somesh and Giri, N.C., "Properties of Tests Concerning Covariance Matrices of Normal Distributions," February 1971, *TR 150*.

Berry, Donald A., "A Bernoulli Two-armed Bandit," February 1971, *TR 151*.

Hwang, Dar-Shong, "Interval Estimation of Functions of Bernoulli Parameters with Reliability and Biomedical Applications," March 1971, *TR 152*.

Nebenzah, Elliott and Sobel, Milton, "Play-the-Winner Sampling for a Fixed-Sample-Binomial Selection," April 1971, *TR 153*.

Sobel, Milton and Berry, Donald A., "A Truncated Inverse Sampling Procedure for Selecting the Best Binomial Population," *TR 154*.

Hoel, David and Sobel, Milton, "New Sequential Procedures for Selecting the Best of k Binomial Populations, With Tables and Comparisons," April 1971, *TR 155*.

Sobel, Milton and Uppuluri, V.R.R., "On Bonferroni-Type Inequalities of the Same Degree for the Probability of a Union," April 1971, *TR 156*.

Thyregod, Poul, "Bayesian Single Sampling Acceptance Plans for Finite Lot Sizes," April 1971, *TR 157*.

Hendrickson and Buehler, "Proper Scores for Probability Forecasters and Determining Subjective Probabilities by Sequential Choices," June 1971, *TR 158*.

Heath and Sudderth, "On a Theorem of De Finetti, Odds Making, and Game Theory," July 1971, *TR 159*.

Lee, J.K. and Sobel, M., "Dorfman and R_1-Type Procedures for a Generalized Group-Testing Problem," September 1971, *TR 160*.

Hwang, Dar-shong and Buehler, Robert, "Confident Intervals for Some Functions of Several Bernoulli Parameters with Reliability Applications," November 1971, *TR 161*.

Das Gupta, Somesh, "Nonsingularity of the Sample Covariance Matrix," April 1971, *TR 162*.

Geisser, Seymour, "A Note on a Linear Discriminant via Idempotent Matrices," November 1971, *TR 163*.

University of Missouri-Columbia, Department of Statistics, 222 Mathematical Sciences Bldg., Columbia, Mo. 65201.

Tsutakawa, R.K., "Design of Experiment for Bioassay," January 1971, *Report No. 42*.

Hewett, John E. and Tsutakawa, R.K., "Two-Stage Chi Square Goodness of Fit Test," June 1971, *Report No. 43*.

(Continued on page 387)