

ESTIMATING THE INTRINSIC RATE OF INCREASE OF A POPULATION FROM THE AVERAGE NUMBERS OF YOUNGER AND OLDER SISTERS

Noreen Goldman

Office of Population Research, Princeton University, Princeton, New Jersey 08540

Abstract—Based on stable population theory, a mathematical relationship is developed between the intrinsic rate of increase (r) of a population and the ratio (Z) of the average number of younger sisters ever born to the average number of older sisters ever born, for a random sample of women in the population. This mathematical formula is then converted into a technique for estimating r from data on numbers of sisters. The extent to which the technique may be generalizable to actual populations is discussed.

INTRODUCTION

Since census estimates and vital registration data are often unreliable or nonexistent, indirect methods for estimating the intrinsic rate of increase (r) of a population have been devised. For example, Coale and Demeny (1967), Bourgeois-Pichat (1958), Keyfitz (1968), and others describe procedures for estimating r from ratios of the number of people in two different age groups of a stable population. In this paper, an alternative indirect technique for estimating r is suggested. The technique utilizes kinship data—specifically, the ratio of the number of younger sisters ever born to the number of older sisters ever born for a random sample of women.

Much of this paper explores, in considerable detail, the theoretical relationship between the ratio of younger to older sisters (Z) and the rate of increase (r) in a stable population. A simple mathematical formula relating r and Z is developed [equation (17)] and is readily converted into an estimation procedure for r [equation (21)]. This procedure appears useful when accurate vital registration data are lacking: data on numbers of sisters can be obtained by survey, and, for some localities, genealogies may already be available

from historical and anthropological sources.

THEORETICAL BACKGROUND

The analytical formulas for the expected numbers of older and younger sisters ever born, for a girl aged a in a stable population, were developed by Goodman et al. (1974, p. 11) and are rederived below. Let us denote the probability of a woman surviving to age x by $l(x)$ and, conditional upon survival to age x , the probability of a woman having a female child between ages x and $x + dx$ by $m(x)dx$, in a one-sex stable population with intrinsic rate of increase r . It is well known that the age distribution of the population is proportional to $e^{-rx}l(x)$. Stated differently, the probability of a woman chosen at random being between ages x and $x + dx$ is proportional to $e^{-rx}l(x)dx$. The age distribution of *mothers* of women in the population is proportional to the product of the age distribution and the age-specific fertility rate $m(x)$. That is, for women of any age, the probability of a woman chosen at random having been born when her mother was between ages x and $x + dx$ is proportional to $e^{-rx}l(x)m(x)dx$, $\alpha \leq x \leq \beta$, where α and β denote the youngest and oldest

ages of childbearing, respectively. Since

$$\int_{\alpha}^{\beta} e^{-rx} l(x) m(x) dx = 1$$

by Lotka's characteristic equation, the proportionality constant equals unity, and the latter probability is exactly equal to $e^{-rx} l(x) m(x) dx$ (Keyfitz, 1977, p. 277).

Let us temporarily impose the condition that a girl aged a (whom we shall call Ego), chosen at random from the stable population, was born when her mother was aged x . Since Ego's mother was alive at age x , she was obviously alive at ages y , $y < x$, and had, on the average, $m(y) dy$ children between the ages of y and $y + dy$, $y < x$. Consequently, if we ignore the infertile periods (pregnancy and postpartum amenorrhea), Ego's mother will have had $\int_{\alpha}^x m(y) dy$ children before the birth of Ego herself, and Ego will have had the same number of older sisters. Furthermore, the chances of Ego's mother surviving for a period of y years beyond the time of Ego's birth equals $l(x+y)/l(x)$, and, conditional upon survival, she will have $m(x+y) dy$ children in the subsequent dy years. Hence, Ego will have had

$$\int_0^a \frac{l(x+y)}{l(x)} m(x+y) dy$$

younger sisters by the time she is aged a .

Recall that we have been conditioning numbers of sisters on Ego's having been born when her mother was age x , an event which has a probability density of $e^{-rx} l(x) m(x) dx$. If we now remove this condition, the expected number of older sisters ever born, $E_a(\emptyset)$, and the expected number of younger sisters ever born, $E_a(Y)$, for a girl aged a are given by

$$E_a(\emptyset) = \int_{\alpha}^{\beta} \left[\int_{\alpha}^x m(y) dy \right] \cdot e^{-rx} l(x) m(x) dx, \quad (1)$$

and

$$E_a(Y) = \int_{\alpha}^{\beta} \left[\int_0^a \frac{l(x+y)}{l(x)} m(x+y) dy \right] \cdot e^{-rx} l(x) m(x) dx, \quad (2)$$

respectively. When we cancel the $l(x)$ in the numerator and denominator, equation (2) immediately reduces to

$$E_a(Y) = \int_{\alpha}^{\beta} \left[\int_0^a l(x+y) m(x+y) dy \right] \cdot e^{-rx} m(x) dx. \quad (3)$$

Note that Ego's age a enters only the expression for younger sisters: since all older sisters are born before Ego herself, their number is independent of Ego's age. When age a is greater than $\beta - \alpha$, the number of younger sisters no longer depends upon a . That is, once a woman is older than $\beta - \alpha$, her mother is beyond the reproductive years (i.e., over age β), and, consequently, the woman has as many sisters as she will ever have.

Equations (1) and (2) for the numbers of older and younger sisters implicitly assume a fertility schedule $m(x)$ applicable to *all* women in the population at all times. More specifically, although we know that Ego's mother had a child at age x , we assume that her past fertility (giving rise to Ego's older sisters) and her subsequent fertility (giving rise to Ego's younger sisters) follow the schedule $m(x)$. We have not taken into account the fact that Ego's mother is certainly not sterile, nor have we allowed for the non-susceptible period following each birth: for example, a girl aged one year or less cannot have any younger sisters, but equation (2) will attribute a small fraction of a sister to her. For a more complete discussion of these problems, see Goodman et al. (1975). The extent to which these simplifications—equivalent to assuming a nonhomogeneous Poisson process of births—may affect the results is described at the end of the paper. (The term "nonhomogeneous" indicates that the birth rate varies with age.)

A THEORETICAL RELATIONSHIP

According to equations (1) and (2), the expected numbers of younger and older sisters are each functions of $l(x)$ and $m(x)$. Since $l(x)$ and $m(x)$ together determine the

rate of increase r (by Lotka's equation), $E_a(Y)$ and $E_a(\emptyset)$ cannot be expressed as functions of r alone. In fact, it has been shown that $E_a(Y)$ and $E_a(\emptyset)$ are much more dependent upon the level of fertility (e.g., the gross reproduction rate) than upon the level of mortality (Goldman, 1977). Thus, the relationship between the expected numbers of sisters and r is an indirect one largely mediated by the level of fertility.

Let us now consider only women past age $\beta - \alpha$ and denote their expected numbers of older and younger sisters by $E(\emptyset)$ and $E(Y)$, respectively (where, as we noted previously, $E(\emptyset)$ and $E(Y)$ are independent of age a). If we look at the ratio Z of the expected number of younger sisters ever born to the expected number of older sisters ever born for these women,

$$Z = \frac{E(Y)}{E(\emptyset)}, \quad a > \beta - \alpha,$$

we discover that Z does have a direct relationship to r . [Note that we want the quantity Z and not Z' ,

$$Z' = E \left(\frac{\text{Number of younger sisters ever born}}{\text{Number of older sisters ever born}} \right) = E \left(\frac{Y}{\emptyset} \right),$$

since Z' calls for division by zero whenever a woman has no older sister.]

The relationship between Z and r can be seen by mathematically manipulating equations (1) and (3). Changing variables in the inner integral of (3) and considering only ages $a > \beta - \alpha$ from now on, we have

$$E(Y) = \int_{\alpha}^{\beta} \left[\int_x^{\beta} l(y)m(y)dy \right] \cdot e^{-rx}m(x)dx. \quad (4)$$

Changing the order of integration in (1) from $\alpha < x < \beta$ and $\alpha < y < x$ to $\alpha < y < \beta$ and $y < x < \beta$, we obtain

$$E(\emptyset) = \int_{\alpha}^{\beta} \left[\int_y^{\beta} e^{-rx}l(x)m(x)dx \right] \cdot m(y)dy. \quad (5)$$

To facilitate comparison between equations (4) and (5), we exchange the x and y variables in (5) and write e^{-ry} as $e^{-rx} \cdot e^{-r(y-x)}$. Equation (5) then becomes

$$E(\emptyset) = \int_{\alpha}^{\beta} \left[\int_x^{\beta} e^{-r(y-x)} l(y)m(y)dy \right] \cdot e^{-rx}m(x)dx. \quad (6)$$

Now, Z , the ratio of the expected number of younger sisters (4) to the expected number of older sisters (6), for a woman of any age greater than $\beta - \alpha$, equals

$$Z = \frac{E(Y)}{E(\emptyset)} = \frac{\int_{\alpha}^{\beta} \left[\int_x^{\beta} l(y)m(y)dy \right] e^{-rx}m(x)dx}{\int_{\alpha}^{\beta} \left[\int_x^{\beta} e^{-r(y-x)} l(y)m(y)dy \right] e^{-rx}m(x)dx}. \quad (7)$$

Equation (7) tells us that $E(Y)$ and $E(\emptyset)$ differ only in the term $e^{-r(y-x)}$ in the inner integral of the denominator of Z . Thus, when the rate of increase equals zero, a woman chosen at random in the stationary population has, on the average, as many older sisters as younger sisters. When r is positive, the term $e^{-r(y-x)}$ is always less than one, and consequently $E(\emptyset)$ is less than $E(Y)$; i.e., in an increasing stable population, a woman chosen at random has more younger sisters than older sisters. Similarly, in a decreasing population, a woman has more older than younger sisters. Thus, (7) tells us the basic nature of the relationship between Z and r but does not express it in a simple functional form.

APPROXIMATING THE RELATIONSHIP

In order to obtain a simple expression relating r to Z , we must in some way

"extract" r from the ratio of integrals in (7). We accomplish this by rewriting (7) in an alternate form [equation (9)] and then applying several approximations to (9).

We begin by using (5) instead of (6) for $E(\emptyset)$ and rewriting (7) as

$$Z = \frac{E(Y)}{E(\emptyset)}$$

$$= \frac{\int_{\alpha}^{\beta} \left[\int_x^{\beta} l(y)m(y)dy \right] e^{-rx}m(x)dx}{\int_{\alpha}^{\beta} \left[\int_x^{\beta} e^{-ry}l(y)m(y)dy \right] m(x)dx} \quad (8)$$

From (8), it follows that

$$\ln Z = \ln E(Y) - \ln E(\emptyset)$$

$$= f(r) - g(r), \quad (9)$$

where $f(r)$ and $g(r)$ are the natural logarithms of the numerator and denominator of (8), respectively. When we linearize $f(r)$ and $g(r)$ by a Taylor series expansion around the point $r = 0$, $\ln Z$ approximately equals

$$\ln Z \approx f(0) + r f'(0)$$

$$- g(0) - r g'(0), \quad (10)$$

where $f'(0)$ and $g'(0)$ are the derivatives of $f(r)$ and $g(r)$, respectively, evaluated at the point $r = 0$. Note that, when $r = 0$, $E(Y) = E(\emptyset)$, and, consequently, $f(0) = g(0)$. Thus, (10) becomes

$$\ln Z \approx r f'(0) - r g'(0),$$

or

$$\frac{\ln Z}{r} \approx f'(0) - g'(0). \quad (11)$$

From equation (8), we obtain $f'(0)$ and $g'(0)$ as given below:

$$f'(0) =$$

$$\frac{- \int_{\alpha}^{\beta} \left[\int_x^{\beta} l(y)m(y)dy \right] x m(x)dx}{\int_{\alpha}^{\beta} \left[\int_x^{\beta} l(y)m(y)dy \right] m(x)dx}, \quad (12)$$

and

$$g'(0) = \frac{- \int_{\alpha}^{\beta} \left[\int_x^{\beta} y l(y)m(y)dy \right] m(x)dx}{\int_{\alpha}^{\beta} \left[\int_x^{\beta} l(y)m(y)dy \right] m(x)dx} \quad (13)$$

Substituting (12) and (13) into (11), we have

$$\frac{\ln Z}{r} \approx \frac{\int_{\alpha}^{\beta} \int_x^{\beta} (y-x) l(y)m(y)m(x)dydx}{\int_{\alpha}^{\beta} \int_x^{\beta} l(y)m(y)m(x)dydx} \quad (14)$$

Since, in general, few deaths occur during the reproductive ages, let us assume that $l(y)$ is approximately constant during this period,

$$l(y) \approx \text{constant} \quad \alpha \leq y \leq \beta. \quad (15)$$

(We will see later that this is a very reasonable approximation.) Cancelling $l(y)$ in the numerator and denominator of (14), we have

$$\frac{\ln Z}{r} \approx \frac{\int_{\alpha}^{\beta} \int_x^{\beta} (y-x) m(y)m(x)dydx}{\int_{\alpha}^{\beta} \int_x^{\beta} m(y)m(x)dydx} \equiv \kappa, \quad (16)$$

where we define κ as the ratio of integrals on the right-hand side of (16). Note that κ is a measure of spread of the fertility schedule $m(x)$. More specifically, κ is the expected absolute age difference between two births (or two siblings) chosen at random from the fertility schedule $m(x)$, if we assume maternal survival of the reproductive period. By replacing $(y-x)$ in the numerator of (16) with $|y-x|$ and integrating y in the numerator and denominator over the entire range from α to β , we can rewrite κ as

$$\kappa \equiv \frac{\int_{\alpha}^{\beta} \int_{\alpha}^{\beta} |y - x| m(y) m(x) dy dx}{\int_{\alpha}^{\beta} \int_{\alpha}^{\beta} m(y) m(x) dy dx}.$$

This formulation for κ shows more clearly that κ is the expected absolute age difference between two births chosen at random from $m(x)$.

Although κ cannot be expressed in terms of the standard deviation σ of $m(x)$, κ and σ are measured in the same units and bear a strong relationship to one another. The relationship between κ and σ can be seen more clearly by replacing the quantity $(y - x)$ with $(y - x)^2$ in the numerator of (16):

$$\frac{\int_{\alpha}^{\beta} \int_{\alpha}^{\beta} (y - x)^2 m(y) m(x) dy dx}{\int_{\alpha}^{\beta} \int_{\alpha}^{\beta} m(y) m(x) dy dx}.$$

Simple algebra reveals that the above ratio equals twice the variance of $m(x)$, $2\sigma^2$. However, since $\sqrt{E[f(x)]} \neq E[\sqrt{f(x)}]$ for general probability density functions $f(x)$, κ cannot be expressed as a constant multiple of σ . Nevertheless, numerical calculations of κ and σ for 49 schedules of fertility reveal that the correlation between κ and σ equals 0.998. Using the standard deviation σ together with skewness yields an almost perfect prediction of κ (multiple correlation of 0.9998).

We have thus accomplished our goal. Rewriting (16), we obtain

$$Z = \frac{E(Y)}{E(\emptyset)} \approx e^{r\kappa}, \quad (17)$$

a very simple expression relating the ratio of younger to older sisters to the rate of increase. Since the derivation of (17) involves two approximations [(10) and (15)], we have yet to determine how well (17) approximates (7). In a later section, we see that, although the approximation in (17) is quite good, we can obtain an even closer one.

A HEURISTIC INTERPRETATION

Equation (17) looks very much like a familiar expression in stable population theory and thus has intuitive appeal. In order to see this, let us denote the number of current births in a stable population by B , and the number of births i and j years ago by B_i and B_j . Then, B_i and B_j equal Be^{-ri} and Be^{-rj} , respectively, and the ratio B_i/B_j equals

$$\begin{aligned} \frac{B_i}{B_j} &= \frac{Be^{-ri}}{Be^{-rj}} \\ &= e^{r(j-i)}. \end{aligned} \quad (18)$$

The resemblance between (17) and (18) is striking. Note, however, that (17) is considerably more complex than (18): since a woman can be a younger and an older sister simultaneously, we are not considering two separate cohorts of women in (17).

For illustrative purposes, we can draw an analogy between numbers of younger and older sisters on the one hand, and the two birth cohorts B_i and B_j on the other. Consider the pool of younger sisters as being births to women at times $\{t_i\}$ and, similarly, the pool of older sisters as being births at times $\{w_j\}$; many women are in both pools and thereby contribute values to both $\{t_i\}$ and $\{w_j\}$. Then, not aiming for rigor, we can replace the ratio B_i/B_j in (18) by $E(Y)/E(\emptyset)$, and the quantity $(j - i)$ by something like the average difference between the values $\{w_j\}$ and the values $\{t_i\}$. We see from (17) that this average difference is, in fact, κ , the average age differential between pairs of siblings in the population.

Equation (17) tells us that the ratio of younger to older sisters in the stable population is jointly determined by the rate of increase r and a measure of spread κ of the fertility curve $m(x)$. For a fixed κ , the higher the r , the larger the ratio of younger to older sisters. Looking at this result intuitively, we note that the younger the age of the mother at which a given woman is born, the more younger sisters

relative to older sisters she can expect to have. Since a large rate of increase of a population results in a younger age distribution of mothers, the larger the r the more younger relative to older sisters a woman chosen *at random* in the population can expect to have.

Equation (17) also tells us that, for a fixed r , the greater the κ the larger is the ratio of younger to older sisters. By taking the derivative of Z with respect to r ,

$$\frac{dZ}{dr} \approx \kappa e^{r\kappa},$$

we see that the larger the κ , the more importance the rate of increase assumes in determining Z . In effect, a larger κ increases the variance of the age distribution of motherhood, thereby allowing r to have more influence on Z . In the absurd limit, a value of zero for κ implies that all sisters are born at a single age of mother and, consequently, that r has no effect on Z .

TESTING THE APPROXIMATIONS

We test the approximations involved in deriving (17) by using typical fertility and mortality schedules. More specifically, for a chosen pair of schedules $l(x)$ and $m(x)$, we evaluate r from Lotka's integral equation $[\int_0^\beta e^{-rx} l(x) m(x) dx = 1]$, Z from equation (7), and κ from equation (16). The integrals are evaluated numerically, using single-year values for $l(x)$ and $m(x)$.

For the analysis described below, 49 fertility patterns were chosen from a set of model fertility schedules (Coale and Trussell, 1974) so as to include all possible combinations of the mean with the standard deviation of $m(x)$. Ten levels of fertility (gross reproduction rates ranging from 0.8 to 3.2) were generated for each fertility pattern, yielding a total of 490 fertility schedules.

In order to test the Taylor series approximation used in (10), we assumed that no deaths occurred in the population [i.e., $l(x) = 1, x \geq 0$]. Using the fertility schedules described above, we obtained 490 values of Z , r , and κ . A regression equation

of the form

$$r = b \left(\frac{\ln Z}{\kappa} \right) \quad (19)$$

was fitted to the set of 490 points. [A regression equation of the form

$$r = c + b \left(\frac{\ln Z}{\kappa} \right)$$

was also fitted to the data points, yielding a value of -0.00004 for c . Since c is so close to zero, we can be confident about omitting a constant term in equation (19).]

Note that, if the approximation in (10) were "perfect," r would be exactly equal to $(\ln Z)/\kappa$ by equation (17), and the value of b would equal one. The fitted regression equation is, in fact, equal to

$$\hat{r} = 0.980 \left(\frac{\ln Z}{\kappa} \right), \quad (20)$$

with an R^2 value of 99.93 percent. The scatter plot appearing in Figure 1 shows the almost perfect fit. (Note that Figure 1 actually displays 490 points, but many points fall in the same locations.) Thus, the Taylor series approximation yields a value for b approximately 2 percent too high. Although (17) has more intuitive appeal than does (20), (20) is obtained empirically and is a better approximation to (7).

We tested the effect of the assumption of no mortality in the reproductive years by determining the regression coefficient b under two alternate levels of mortality: $e_0 = 30$ years and $e_0 = 50$ years. [Life tables were chosen from the Coale and Demeny West series (Coale and Demeny, 1966).] Both life tables, of course, involve some deaths during the reproductive years. The new regression coefficients and measures of goodness of fit appear in Table 1. Since b is virtually invariant with different $l(x)$ schedules, the approximation in (15) is justifiable. In summary, we can accept (20) as an "almost perfect" approximation to (7) under all levels of mortality.

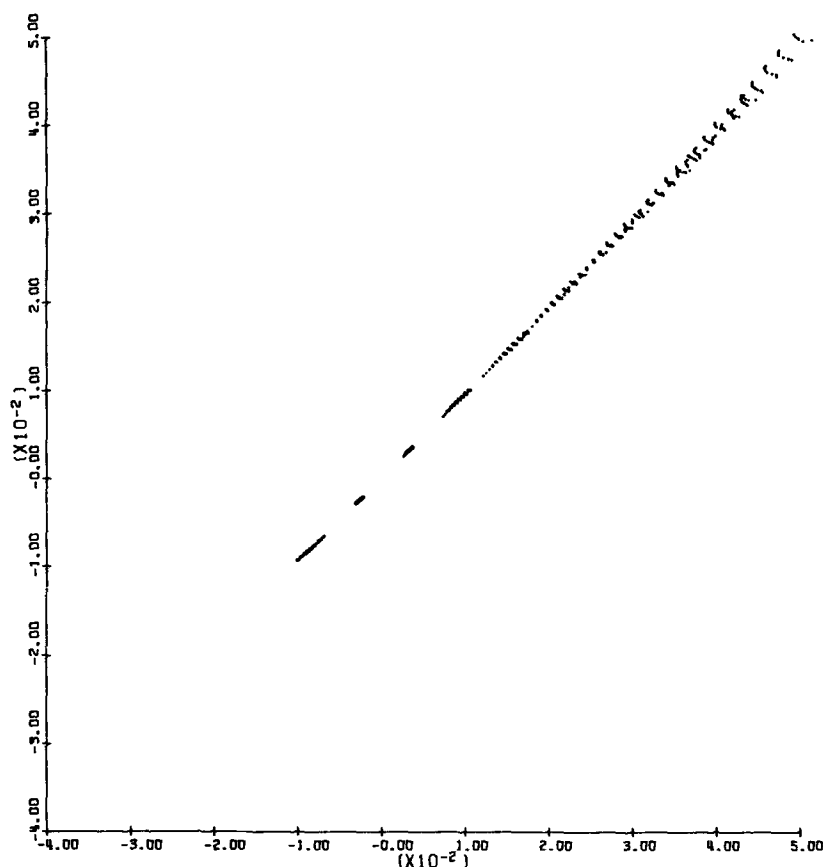


Figure 1.—Scattergram of $\ln Z/\kappa$ Versus r for 490 Fertility Schedules Under the Assumption of No Mortality

AN ESTIMATION PROCEDURE FOR r

In addition to its extreme simplicity, equation (20) appears to provide a straightforward estimation procedure for r when estimates of Z are available from censuses, sample surveys, or genealogies. However, in order to apply (20), we must secure an estimate of κ , a quantity not frequently used as a statistical measure of spread. The high correlation between κ and σ suggests that we can obtain an analogue to equation (20) by replacing κ with σ .

Using calculated values of σ for the 49 fertility schedules, we fit the regression equation

$$r = b' \left(\frac{\ln Z}{\sigma} \right)$$

to the data points for the three levels of mortality. The results, appearing in Table 2, reveal that the equation

$$\hat{r} = 0.921 \left(\frac{\ln Z}{\sigma} \right) \quad (21)$$

fits the data points almost as well as does (20) and offers us an alternative approximation to (7). Since values of σ have frequently been computed for fertility schedules (e.g., Coale and Trussell schedules), equation (21) provides a more practical estimation procedure for r .

ASSUMPTIONS OF THE MODEL

Until now, we have not considered in detail the assumptions which underlie equations (1) and (2). These assumptions, of course, are also incorporated into the

Table 1.—Regression Coefficients b and R^2 Values (Percent of Variance Explained) for Equations of the Form $r = b \ln Z/\kappa$

| Level of Mortality | b | R^2 |
|--------------------|-------|-------|
| $e_0 = 30$ | 0.972 | 99.93 |
| $e_0 = 50$ | 0.984 | 99.94 |
| No mortality | 0.980 | 99.93 |

estimation procedures [equations (20) and (21)] and thus determine their applicability to actual populations. The three major assumptions of our model are as follows.

1. Fertility and mortality are unchanging with time, so that we have a stable age distribution in the population.
2. Births occur as a nonhomogeneous Poisson process. That is, the maternity function for *all* women aged x is taken to be $m(x)$, regardless of whether some women are known to be fertile or have just had a child prior to age x (Goodman et al., 1975).
3. Fertility and mortality act independently of one another: i.e., those women who die before age $\beta - \alpha$ and cannot be sampled would eventually have the same number of sisters as do those women who survive beyond $\beta - \alpha$.

Whereas the third assumption may be largely true, assumptions (1) and (2) will almost surely be violated in any real popu-

Table 2.—Regression Coefficients b' and R^2 Values (Percent of Variance Explained) for Equations of the Form $r = b' \ln Z/\sigma$

| Level of Mortality | b' | R^2 |
|--------------------|-------|-------|
| $e_0 = 30$ | 0.913 | 99.96 |
| $e_0 = 50$ | 0.926 | 99.91 |
| No mortality | 0.921 | 99.84 |

lation. The extent to which our model is robust to these assumptions has yet to be fully tested, but some preliminary remarks can be made.

It appears that, in fact, the relationship between Z and r is sensitive to departures from a stable age distribution. We can see this phenomenon more easily by considering women of a particular age x at time t . Loosely speaking, older sisters are births in the five years prior to time $(t-x)$, whereas younger sisters are births in the five years subsequent to $(t-x)$. If the birth rate has been changing over this ten-year time period, we expect Z to be sensitive to these fluctuations in fertility. It is essential in such a nonstable population to stratify the ages of women in the sample, thereby obtaining separate estimates of Z and r for women in different age groups. Each value of r no longer represents an "intrinsic" rate of increase, but rather it measures the rate of increase of births during a period of time centered about the birth date of the women sampled.

We can partly relax the assumption of a nonhomogeneous Poisson process of births without affecting the model. Specifically, we can replace the $m(x)$ in the inner integrals of (1) and (2) with $Km(x)$, where K is constant, without altering (7) or any subsequent equations. By introducing K , we allow for the fact that women who have already had a birth (the girl in our sample) cannot be sterile. Such women are likely to have a fertility curve higher than $m(x)$, so that K will generally be greater than one. Although the model also ignores the nonsusceptible period following conception, this simplification affects mostly the numbers of sisters to very young girls, not the numbers to women older than $\beta - \alpha$. The extent to which departures from the Poisson process assumption affect the model can be further assessed by simulation methods.

CONCLUSIONS

The relationships between Z and r given in equations (7), (20), and (21) are theoretical ones. To a large extent, they can be

viewed as consequences of stable population theory. In order to determine how well these relationships hold up in real populations, the model would have to be tested on actual data: the sensitivity of the model to the assumptions of a non-homogeneous Poisson process of births and of a stable age distribution could thereby be assessed. More realistically, the model would have to be redefined somewhat in order to deal with nonstable populations.

Should the relationship between Z and r be generalizable to real populations, we have a very straightforward technique for estimating r in the absence of vital data. Numbers of sisters can be ascertained by survey: questions on the numbers of younger and older sisters ever born can easily be incorporated into existing censuses and surveys. Respondents who do not even know their own ages are likely to know the "ordering" of their siblings. (Unfortunately, however, respondents are less likely to know about older siblings who died before they themselves were born.) For some localities, genealogies may already be available from historical and anthropological sources. Having obtained data on siblings, we need only divide the average number of younger sisters ever born to women in a sample by the average number of older sisters ever born to these women and substitute this estimate of Z along with an estimate of σ (obtained from some knowledge of the age distribution of fertility) into equation (21). As long as women in the sample are over age $\beta - \alpha$, the ages of women can be ignored in the estimation procedure.

ACKNOWLEDGMENTS

I would like to thank Nathan Keyfitz, Ansley Coale, Douglas Ewbank, Bryan Boulier, James Trussell, Jane Menken,

Mike Stoto, and Gary Littman, all of whom offered valuable suggestions on various drafts of the paper. I am also indebted to an anonymous referee who suggested the analytical proof for deriving equation (16) from equation (7). The Fortran program used for the numerical integrations was adapted from programs written by Thomas Pullum and James Trussell. Much of this work was undertaken as part of my doctoral dissertation at Harvard University, under financial assistance from the Harvard School of Public Health.

REFERENCES

- Bourgeois-Pichat, Jean. 1958. Utilisation de la Notion de Population Stable Pour Mesurer la Mortalité et la Fécondité des Populations des Pays Sous-Développés. *Bulletin de l'Institut International de Statistique* (Stockholm, Actes de la 30^e Session) 36:94-121.
- Coale, Ansley J., and P. Demeny. 1966. *Regional Model Life Tables and Stable Populations*. Princeton: Princeton University Press.
- , and P. Demeny. 1967. *Methods of Estimating Basic Demographic Measures From Incomplete Data. Manuals on Methods of Estimating Population, Manual 4. ST/SOA/Series A/42*. New York: United Nations.
- , and T. J. Trussell. 1974. Model Fertility Schedules: Variations in the Age Structure of Childbearing in Human Populations. *Population Index* 40:185-258.
- Goldman, Noreen. 1977. *The Demography of Kin*. Unpublished D.Sc. dissertation. Cambridge, Mass.: Department of Population Sciences, Harvard University.
- Goodman, L. A., N. Keyfitz, and T. W. Pullum. 1974. Family Formation and the Frequency of Various Kinship Relationships. *Theoretical Population Biology* 5:1-27.
- , N. Keyfitz, and T. W. Pullum. 1975. Addendum to "Family Formation and the Frequency of Various Kinship Relationships." *Theoretical Population Biology* 8:376-381.
- Keyfitz, Nathan. 1968. *Introduction to the Mathematics of Population*. Reading, Mass.: Addison-Wesley Publishing Company.
- . 1977. *Applied Mathematical Demography*. New York: John Wiley and Sons.