Chapter VII

EXAMPLES OF ESTIMATES BASED ON QUESTIONS ABOUT FERTILITY AND MORTALITY

A. Estimation of fertility from reports on child-bearing in censuses or surveys

The estimation of fertility from data on (*a*) births in the year before a survey or census, and (*b*) children ever born, is illustrated by an example based on hypothetical data. A synthetic example is used because the only instances of surveys containing the requisite information are in tropical Africa, and, as is noted in chapter IV, the validity of the method is sometimes seriously impaired by the extensive age-misreporting characteristic of surveys and censuses in Africa. The suitability of the method for estimating fertility in an African population must be decided after a detailed examination of the quality of the data in the survey in question—an examination that would go beyond the scope of this manual. The reader is referred to the detailed studies of different populations in the book on the population of Africa recently completed at the Office of Population Research.[1]

The method exemplified in table 33 is of greatest potential value for populations (such as in Latin America, the Philippines, and especially in the Republic of Korea and Thailand) for which age-misreporting is less extreme. The example shows how data from such areas can be employed when the appropriate questions have been included in a census or survey.

[1] Brass, *et. al.*, op. cit.

Total fertilility derived from the age specific fertility rates reported in the example (column 3) is

$$5 \sum_{i=1}^{7} f_i ,$$

or 5.24. The adjustment of these fertility rates for a possible error in the reference period is achieved by calculating the average value of cumulative fertility in the age intervals shown in column 2 (F_i in column 7), and forming the ratio of reported parity (P_i) to cumulated fertility (F_i). Reported parity at 20-25 (P_2) and 25-30 (P_3) is assumed to be approximately correct, although the latter may be affected to some slight degree by omissions. The ratio of P_2/F_2 is a correction factor that makes the fertility rates consistent with the average number of children ever born reported by women 20-25. In column 9 there are age specific fertility rates that have been multiplied by P_2/F_2 (1.313). The adjusted estimate of total fertility is five times the sum of the rates in column 9, or 1.313 times the figure derived from the unadjusted fertility rates. The estimate is 6.88—higher than either the average parity reported by older women (deficient because of omissions) or than the cumulation of reported fertility (deficient because of a shortened reference period).

Computational procedure. The fertility rates in column 3, based on births reported for the year before the survey, pertain to women one-half year younger than the intervals in column 2. Therefore the estimation of average cumu-

Table 33. The estimation of total fertility as P_2/F_2 times reported total fertility

Interval (i) (1)	Exact age of woman at time of survey (2)	Average number of births in preceding year per woman (f_i)[a] (3)	Average number of children ever born P_i (4)	Cumulative fertility at beginning of interval $\left(\sum_{j=0}^{i-1} f_j\right)$ (5)	Multiplying factors for estimating average value fertility (w_i) (6)	Estimated average cumulative fertility $\left(F_i = \sum_{j=0}^{i-1} f_j + w_i f_i\right)$ (7)	P_i/F_i (8)	Adjusted age specific fertility rates $(f'_i = f_i \times P_2/F_2)$ (9)
1	15-20	.081	.186	0	1.963	.159	1.170	.106
2	20-25	.242	1.435	.405	2.842	1.093	1.313	.318
3	25-30	.261	3.109	1.615	3.011	2.401	1.295	.343
4	30-35	.238	4.176	2.920	3.121	3.663	1.140	.312
5	35-40	.166	4.710	4.110	3.247	4.650	1.013	.218
6	40-45	.043	4.761	4.940	3.548	5.093	.935	.056
7	45-50	.017	4.503	5.155	4.484	5.231	.861	.022

[a] For age intervals one-half year less than shown in column 2.

lative fertility is obtained by the use of multiplying factors found by interpolation in annex table IV.1. The key to the interpolation is f_1/f_2, equal to .335. The multiplying factors in column 6 were (5/130) × entries where f_1/f_2 equals .460 plus (125/130) × entries where f_1/f_2 equals .330.

The entries in column 8 are typical of those found in African surveys least distorted by age-misreporting. P_1/F_1 is best ignored because of the intrinsic difficulty of estimating F_1 . P_2/F_2 and P_3/F_3 are approximately equal, and either ratio can be taken as a multiplying factor for the correction of fertility rates for bias in the reference period. The steady decline in P_i/F_i past age thirty is the typical result of progressively greater omission of children ever born by older women.

B. Estimation of mortality from reported numbers of children ever born, and children surviving

Preliminary adjustment of data. This method of estimation requires responses from a census or survey on the number of children ever born alive to each woman, and the number of children surviving, with average numbers per woman tabulated for the standard five-year age intervals. Deficiencies in the data for which adjustments can sometimes be made include: (*a*) a moderate proportion of women for whom no responses are obtained; (*b*) asking the relevant questions (or tabulating responses) only for married women, or only for women ("mothers") who have experienced at least one live birth.

(a) *Adjustment for non-response.* The women for whom children ever born are *not* tabulated are not, ordinarily, representative in their average parity of the age group to which they belong. There is a widespread tendency to leave a blank space instead of a zero for the response to the question of the number of children ever born for women of zero parity. The evidence of this tendency is a strong positive correlation between the proportion of women reported childless in each age interval with the proportion of non-responses. If all non-responses were of this sort, they should be counted as zeros in calculating the average number of children ever born in each age interval. However some of the non-responses represent a genuine absence of information, so that the assumption of zero parity would produce a downward-biased estimate. El Badry has proposed[2] a simple but often effective technique for determining approximately which non-responses represent zero parity and which the genuine absence of information (as, for example, when data are supplied by a neighbour). He suggests fitting a straight line of the form $y = ax+b$, where the observed values of y are the proportion of non-respondents in each age interval and x is the proportion reporting zero parity. If the observed relationship is very closely fitted by the straight line, it may plausibly be assumed that the fraction of non-responses genuinely associated with the absence of information is b, and that the proportion of childless women recorded as non-respondents is $a/1+a$. The recommended method of adjusting the non-responses, then, is to make a scatter diagram showing the proportion of non-responses in each age interval on one axis, and the proportion of childless women on the other. If the resultant points are closely fitted by a straight line, extend this line to the zero value of the proportion childless. The proportion of non-responses on the straight line at this point (when the proportion childless is hypothetically zero) can be taken as an estimate of the true proportion of non-responses. The recommended arithmetical adjustment of the data on the average number of children ever born and on the average number surviving is the omission from the denominator of the estimated number of true non-respondents, and the inclusion in the denominator of the estimated number of non-respondents who are considered to be in fact childless. In most instances almost all of the non-respondents fall in the probably childless category, and little or no error is introduced by assuming that all non-responses indicate zero children.

(b) *Adjustments when data are tabulated only for married women or for mothers.* In some censuses and surveys, questions about children ever born are asked (or at least tabulated) only for ever married women, and in others, data are shown only for women who have borne at least one child. A word about the nature of the resulting biases is not out of place. First, if illegitimacy is infrequent, the proportions surviving among children ever born will be adequately representative, and the only effect of obtaining information solely from ever married women is that it is necessary to estimate the average parity of *all* women in each age group indirectly, usually on the assumption that all births occur to married women. If, in fact, illegitimate births contribute significantly to the average parity of women 15-19, and not so much to women 20-24, the ratio P_1/P_2 may be underestimated, leading to the selection of adjustment factors in table V.1 that tend tooverestimate $q(2)$ and $q(3)$. The bias resulting from the higher mortality rates experienced by children born to non-married women, on the other hand, tends to cause an underestimate of mortality, although this latter bias is minimized by the fact that women under twenty bearing children while single are often married by the time (at age 20-24 or 25-29) their fertility histories are used to estimate ${}_2q_0$ and ${}_3q_0$.

The adjustment made when information is given only for ever married women is to determine average parity in the two relevant intervals (15-19 and 20-24) by multiplying the average parity of ever married women by the ratio ever married women/total women. This adjustment cannot be accepted as valid in populations where a major proportion (say more than 10 per cent) of the births to women 15-19 occur to the non-married. When data are supplied only for "mothers", the adjustment is analogous to that for ever married women, i.e., parity is estimated by multiplying the average parity of mothers by the ratio mothers/total women. In both instances of limited data, the reported proportion surviving among children ever born is accepted as representative of the experience of all women. In fact the proportion surviving for ever married women is higher than for the non-married, and this bias of course also holds for "mothers" if the

[2] M.A. El Badry, "Failure of enumerators to make entry of zero: errors in recording childless cases in population censuses", *Journal of the American Statistical Association*, vol. 56, No. 296, December 1961, pp. 909-924.

TABLE 34. CALCULATION OF $_1q_0$, $_2q_0$, $_3q_0$, $_5q_0$, $_{10}q_0$, $_{15}q_0$, AND $_{20}q_0$ FOR BRAZIL, BASED ON CHILDREN EVER BORN, AND CHILDREN SURVIVING RECORDED IN THE 1950 CENSUS

Interval (i) (1)	*Age of women* (2)	*Average number of children ever born* (P_i) (3)	*Average number of children surviving* (S_i) (4)	$1-S_i/P_i$ (5)	*Multipliers for column 5 from* P_1/P_2 (6)	*Age* x (7)	*Proportion dead by age* x ($_xq_0$) (8)
1	15-19	.146	.118	.1918	1.058	1	.203
2	20-24	1.099	.870	.2083	1.050	2	.219
3	25-29	2.516	1.947	.2262	1.016	3	.230
4	30-34	3.883	2.935	.2442	1.019	5	.249
5	35-39	5.065	3.730	.2636	1.029	10	.271
6	40-44	5.778	4.146	.2825	1.007	15	.284
7	45-49	6.212	4.353	.2993	1.006	20	.301

q. 1.30 – .54 (1.099/2.516) x .870.

questions have been asked only of married women, and tabulated only for those with at least one child. If, in fact, the questions are *asked* only of women who have borne at least one child, the probable bias is increased because of the likelihood that "mothers" with no surviving children may have been excluded.

Computational procedure. The estimation of mortality from child survival rates is illustrated in table 34 on the basis of data taken from the 1950 census of Brazil. Columns 3 and 4 show the average number of children ever born per woman and the average number surviving for the age intervals shown in column 2. The proportion of non-survivors is shown in column 5; these proportions are converted into estimates of $_1q_0$, $_2q_0$, ..., $_{20}q_0$ by multipliers taken from annex table V.1. P_1/P_2 is .146/1.099 = .133. The multipliers in column 6 were obtained by interpolating between the columns in table V.1, for which P_1/P_2 is .143 and .090, specifically, by adding (.811) × (entries where P_1/P_2 = .143) and (.189) × (entries where P_1/P_2 = .090). The final estimates are in column 8.

Comments. As is pointed out in the discussion in section B of chapter II the level of infant mortality derived from child survival reported by women 15-19 should not be regarded seriously, because of the basic weakness of the method of estimation at this point. A better estimate of $_1q_0$ is obtained by accepting $_2q_0$, and the relationship between $_1q_0$ and $_2q_0$ in the model tables. The resultant estimate of infant mortality is .171, if based on the "West" female model tables, and .178 if based on the males, so that an acceptable estimate of $_1q_0$ is about .175.

The level of mortality indicated by the sequence of estimates of $_xq_0$ implies the following sequence of $^0e_0$'s for both sexes (as calculated from annex tables I.2 and I.1):

$_xq_0$	*Corresponding* $^0e_0$
$_2q_0$	41.2
$_3q_0$	42.2
$_5q_0$	42.4
$_{10}q_0$	42.4
$_{15}q_0$	42.6
$_{20}q_0$	42.9

The remarkably consistent sequence of the implied $^0e_0$ values (or of the implied mortality levels in general) suggests that the basic data are of very good quality. This is also supported by an examination of the reported parity distribution itself which shows no obvious signs of an increasing failure to report children as the age of the reporting women progresses. Naturally these observations inspire increased confidence in the mortality estimates derived above. On the other hand, it should be borne in mind that survival rates to increasingly higher ages reflect the mortality experience of an increasingly longer period prior to 1950. It is possible, therefore, that there is a fortuitous element in the high consistency of the mortality levels implied by these rates; namely, the effects of the actually higher mortality of earlier periods to some extent might have been offset by the probable tendency on the part of older women to omit a higher proportion of their children who are dead than of those who are still alive.