The Own-Children Method of Estimating Age-Specific Fertility Rates: Lecture Notes Prepared for a Demographic Analysis and Data Evaluation Workshop, Department of Statistics, Kuala Lumpur, Malaysia, 10-28 March, 1975

APRIL 1975

Griffith Feeney
The East-West Center
Honolulu, Hawaii
96822

CONTENTS

the United States (Differential Current Fertility in the United States, by Lee-Jay Cho, Wilson H. Grabill, and Donald J. Bogue. Chicago: Community and Family Study Center, University of Chicago, 1970). Cho subsequently began working in Korea snd Malaysia and, faced with a lack of vital registration data, generalized the method so as to allow the estimation of annual fertility statistics for the ten year period preceding a given census. The necessary tabulations were produced for the first time from the 1966 Korean census under Cho's guidance (Korean Office of Statistics and Research Seoul, Korea, The 1966 Census of Population A Special Tabulation Report, Fertility: Women by Number of Own-Children, August 1970). The method was refined and extended in a number of directions by Cho during the late 1960's and plans for application to the 1970 round of censuses in several Asian countries were made. Since 1970 further developments and analysis of the resulting special tabulation has been carried out by Cho and co-workers at the East-West Population Institute in Honolulu, Hawaii. A report on the application of the technique to Korea has just been published, Estimates of Current Fertility for the Republic of Korea and its Geographical Subdivisions: 1959-1970 (Seoul: Yonsei University Press 1975). Both publications include a general discussion of the method as well as the application to the particular countries.

3. ADVANTAGES AND LIMITATIONS OF THE METHOD

Advantages of the method are: it provides extremely detailed estimates, age-specific fertility rates by single years of age for each of fifteen years prior to the census; only one census is required; the population need not be closed to migration; the method is relatively insensitive to recall errors. Limitations of the method are: it requires a special census tabulation; it is sensitive to census undercounting and misreporting errors (these may be adjusted for providing a post-enumerative survey or other census evaluation is available); it requires a considerable volume of calculation.

4. ESTIMATION OF BIRTHS BY REVERSE SURVIVAL

Children aged 0 in completed years at the time of a census were necessarily born during the one year period preceding the census, children aged 1 in the second year preceding the census, and so forth. By inflating the numbers of children at each age to allow for mortality one obtains estimates of births during each year preceding the census. Note that the word "year" here refers to one year periods beginning and ending at the month and

day of the census, not to calender years.

This procedure for estimating births from the census age distribution is well known and is often applied to test the consistancy of census and birth registration data. It requires a life table with L_x values given by single years of age. For comparisons with birth registration data it is usually necessary to make an adjustment for the difference between calender years and years beginning and ending on the month and day of the census. Table 1 shows this method applied to Peninsular Malaysia using 1970 census data and birth registration data for 1966-1970. The results suggest an underenumeration of children 0-4 years old on the order of 5 percent, a reasonable figure. The true figure is probably somewhat higher due to some underregistration of births. The derivation of the reverse survival factors is explained below.

5. ESTIMATION OF BIRTHS BY OWN-CHILDREN METHOD

The innovation in the own-children method is the observation that, given a suitably constructed census tabulation, these estimated numbers of births may be distributed by age of mother. The special census tabulation required consists of all persons recorded in the census as being under 15 years of age classified by single year of age and single year of age of mother. Single years of age of mother should normally be tabulated from 10 to 64 and not for any smaller range.

A complication enters in connection with the classification by age of mother. Birth registration statistics normally classify births in a given year by age of mother in completed years at the time of birth. Own-children tables, by contrast, give the distribution of births by age of mother as of the beginning of the year of birth. It would perhaps not be amiss to reiterate here that "year" in the own-children context means a one year period beginning and ending at the time of the census. This classification by age of mother at the beginning of the year instead of at the time of birth is in fact inevitable. The age of the mother in completed years (if known) is known as of the time of the census, and, by subtraction, one year prior to the census, two years prior to the census, and so on, but age in completed years as of any other time is indeterminate. Consider for example a woman aged exactly 27.5 years at the time of the census with a child aged 0 in completed years. If the age of the child is 3 months, the age of the woman in completed years at birth would have been 27; but if the child is 9 months old, the mother's age would have been 26. One may conclude that the

mother's age in completed years at birth was either 26 or 27, but it is impossible to choose between these two ages. This ambiguity is due to knowing age only in completed years and not exact age.

This dilemma may be resolved by assuming that, of the total births occurring during a given year to women aged x in completed years at the beginning of the year, one half occurred to mothers aged x in completed years and one half to mothers aged x+1 in completed years. This is analogous to the conventional assumption of a separation factor equal to one half. The number of births during a given year to women aged x in completed years at birth may then be estimated as the average of births to woman aged x-1 in completed years at the beginning of the year of birth and women aged x in completed years at the beginning of the year of birth.

6. ADJUSTMENTS FOR CENSUS UNDERENUMERATION AND AGE MISREPORTING

It has already been noted that the own-children method is sensitive to census underenumeration and age misreporting errors. The reason is simple enough. Since births are estimated by inflating census counts of children at each single year of age, and underenumeration of children at a given age will result in an underestimate of births a corresponding number of years prior to the census. Likewise, misreporting of children's age will result in incorrect distribution of estimated births among the various years preceding the census. underenumeration errors on the order of five percent would not be unexpected even in a census well conducted in favorable circumstances and might rise to ten percent or more for particular population subgroups or in less favorable circumstances. There are of course populations for which available data is such that errors of this magnitude may be fairly tolerable. In most circumstances, however, it will probably be desirable to undertake an evaluation of census under-enumeration and age misreporting errors and to calculate on this basis a set of adjustment factors to reduce the effect of census errors. Such an evaluation is generally undertaken for principle population subgroups in any case and need not pose much additional burden.

It should be emphasized, however, that adjustment factors calculated for the population at large will not in general be applicable to the populations of geographical subdivisions or to subpopulations defined by various socioeconomic characteristics. Thus, for example, extremely low income or low education groups may suffer considerably greater under-enumeration and age

misreporting than the population at large. If analysis of fertility differentials is contemplated, therefore, the additional task of census evaluation for the relevant population subgroups should be undertaken. Such analysis, not normally undertaken as part of post-census operations, involves a labor proportional to the detail of the fertility differentials to be studied.

The example of own-children estimation given below does not include any analysis or adjustment for census underenumeration or age misreporting errors. This was inevitable by reason both of the time available for this presentation and by the absence of any census evaluation analysis at the district level. The crude calculation in Table 1 suggest an underenumeration of about five percent, but there is no basis for supposing this figure applicable to the particular administrative district at Selangor for which the example has been carried out. Were adjustment factors available, they would simply be included along with the factors for mortality and the nonown-children adjustment. All other details of the example given in the following pages would be unaffected.

Although inevitable, the omission of census evaluation analysis and calculation of appropriate adjustment factors is in one respect regrettable, for it will usually be essential to the application of the own-children method in practice. It is also true, however, that the nature of census evaluation analysis will vary from one census to another according to the customs and circumstances of the population, the availability of vital statistics or post-enumerative survey data, and so forth. By contrast, the tabulations and procedure for application of the own-children method apart from the derivation of these adjustment factors are virtually the same for all populations. There is accordingly some logical justification for considering the own-children estimation technique and the census evaluation analysis frequently necessary for its application under separate headings.

7. PROCEDURE FOR DETERMINING AGE OF MOTHER

Age of mother is determined from the "current age" and "relationship to head of household" questions on the census form. An illustrative procedure is given in Figure 1. There will be erroneous age determination where adopted children are recorded as sons and daughters of the head of the household, and where the household includes sons and daughters of a male head of household outside the current marriage. It is reasonable to

expect, however, that these errors will be minimal in most situations.

8. DETERMINATION OF REVERSE SURVIVAL RATIOS FOR CHILDREN

Life tables are rarely given for both sexes combined, hence L_x values for reverse survival of children must usually be calculated as averages of male and female values weighted according to the sex ratio at birth. Available life tables will not usually provide single years of age, hence some type of interpolation procedure is necessary. For ages 5 and over any of the conventional techniques will usually give satisfactory results, even simple linear interpolation on l_x values for x = 5, 10, ... These methods will frequently give poor results under age 5, however, and it may be preferable to revert to the Coale–Demeny model life table families to obtain values for x = 2, 3, 4. The procedure is exemplified in Table 3 and explained in the notes to this table. A similar procedure may be used to determine model life table values corresponding to Brass estimates of q_0 based on child survivorship. Once values are obtained for each sex separately, the l_x applicable to both sexes together must be calculated by taking a weighted average based on the sex ratio at birth. Values of L_x for both sexes may then be calculated and the desired reverse survival ratios obtained as $1/L_x$. The entire procedure is exemplified in Table 4 and described in the notes to this table.

9. ESTIMATION OF FEMALE AGE DISTRIBUTIONS PRIOR TO THE CENSUS

The first step is to determine life table L_x values for females by single years at age for x = 10, 11, ..., 64. Given an abridged life table, any of the various techniques for splitting groups into fifths may be applied directly to the ${}_5L_x$ values, x = 5, 10, ..., 65. Alternatively, L_x values for single years may be obtained by any of various interpolation techniques and L_x calculated as $0.5x(l_x + l_{x+1})$. The latter procedure is exemplified in Table 5 with simple linear interpolation.

The estimated female age distributions one, two, ..., years prior to the census are most efficiently obtained by calculating, for each age x = 11, 12, ..., 64, the census count of women aged x divided by Lx, storing this value in the calculator, and multiplying in turn by L_{x+1}/L_x. This yields, respectively, estimates of women aged x–1 one year prior to the census, x–2 two years prior to the census, and so on. This corresponds to following the cohort of women aged x at the time of the census backward in time. The age distributions in Table 6

were estimated in this manner using the L_x values given in Table 5.

10. PROCEDURE FOR CALCULATING OWN-CHILDREN ESTIMATES OF FERTILITY

The first step is of course to create the special own-children tabulation from the original census schedules. This will always be done by machine, except perhaps for test tabulations for very small population subgroups. The distribution of women aged 10-64 years by single years of age is also required, but this is almost universally tabulated in any case. All the necessary data may be collected in a single table, exemplified in Table 2. The data in Table 2 refer to Kuala Langat, an administrative district in Selangor, and were obtained from the Department of Statistics worksheet shown in Attachment 1. They were used to illustrate the application of the method. Numbers of children aged 10-14 are not available.

The second step is to obtain male and female life tables applicable to the population during the fifteen years preceding the census. It is possible to utilize several pairs of life tables for different periods to allow for changing mortality, but the quantitative effect on the fertility estimates will usually be too small to justify the extra effort required. For the example worked out here the 1970 life tables for Peninsular Malaysia published by the Department of Statistics were used.

The third step is to perform a census evaluation analysis for under enumeration and age misreporting of children in each population subgroup f or which estimates are to be made, and to calculate on this basis adjustment factors to be applied to children at each single year of age.

With these basic data in hand the preliminary computational steps are as follows.

1. Calculate single year L_x values for females, x = 10, 11, ..., 64. This is shown in Table 5.

2. Calculate the age distributions of females prior to the census by reverse survival of the census age distribution. These age distributions are shown in Table 6.

3. Make a table of own-children classified by year of birth and age of mother at beginning of

year of birth. This is exemplified in Table 7. This table is nothing more than a copy of the basic own-children tabulation (Table 2) in a slightly different format and is not required. It is nonetheless advisable when working with a desk calculator to avoid careless errors.

4. Calculate nonown—children adjustment factors for each age x = 0, 1, ..., by dividing the total number of children age x by the number of own—children age x. Enter these factors at the top of the table which will contain the final estimates of age—specific fertility rate, exemplified in Table 8.

5. Calculate reverse survival ratios for children. This is shown in Tables 3 and 4.

6. Calculate net adjustment factors as the product of the factors 4 and 5 above and the census error adjustment factors, if used.

The final step is to calculate the age—specific fertility rates, working from the age distribution table (Table 6) and the rearranged own— children table (Table 7). For each year, calculate the age—specific fertility rate for age x as follows.

(i) Calculate the average of (a) the number of own—children born in this year to women aged x at the beginning of the year and (b) the number of own—children born in this year to women aged x+1 at the beginning of the year. For the first year preceding the census and x 27 years, for example, calculate (154+165)/2 = 159.5, the figures in parentheses coming from the last column of Table 7 opposite ages 26 and 27.

(ii) Multiply this average by the net adjustment factor for the year. This yields an estimate of the number of births during the year to women aged x at birth. The net adjustment factor for the first year preceding the census is 1.10583 (Table 8, third row, last column), hence the estimated number of births Exception: When x is the youngest age for which own—children are tabulated, use the number of own-children to women aged x plus one-half the number aged x+1 in the year preceding the census to women aged

27 at birth is 174.5 x 1.10583 =192.96....

(iii) Divide this estimate of births by the average of the number of women aged x at the beginning of the year and the number aged x at the end of the year. The latter average is an estimate of the mean number of women aged x during the year, and the quotient is the age-specific fertility rate for women aged x during the given year. Continuing the example begun above, one divides 192.96... by (562 + 654)/2 = 608, which yields 0.3174 as the age-specific fertility rate for women aged 27 in the year preceding the census.

Note that, as a computational shortcut, the division by 2 in steps (ii) and (iii) may be omitted, as this division occurs in both numerator and denominator of the calculated rate.

11. ANALYSIS OF THE RESULTS

The age-specific fertility rates for the first year preceding the census are plotted in Figure 2. Their sum over ages 15-54 gives an estimated total fertility rate of 5.92 children per women. We proceed to compare these results with rates calculated from birth registration data.

The pattern of the age-specific fertility rates in Figure 2 shows a number of minor anomalies which are probably due to errors in the source data. Since the age pattern of the rates reflects the census female age distribution, one looks to age misreporting errors to explain these anomalies. This histogram of the census age distribution shown in Figure 3 suggests considerable age misreporting, but it remains to explain how this misreporting might explain the irregularities in the age-specific fertility rates shown in Figure 2. This is an area in which very little research has been conducted and, accordingly, no conclusions of substance can be drawn. Several general observations may be made, however. First, the effect of age misreporting on the own-children estimates is confounded by the averaging procedure which occurs in both the numerator and the denominator of the estimated rates. Second, since women and their children are linked together in the estimation procedure, transference of a woman's age classification results in a corresponding transference of her children into the appropriate "age of mother" category. The effect of age misreporting is thus to give a fertility rate for each age which is in fact a sort of

weighted average of rates of or this and surrounding ages.

Registered births in Malaysia in 1970 were tabulated by age and state of residence of mother as well as by age of mother and state of occurrence of birth. Tabulations of births in administrative districts within states are available, but by district of occurrence only and without the age of mother classification. We cannot, therefore, calculate age-specific birth rates for Kuala Langat directly from the birth registration data. Table 9 shows age-specific birth rates for Clangor state (which includes the administrative district of Kuala Lang at) as a whole, however, and we see at once that the total fertility rate calculated from the birth registration data for the state is far lower than the own-children estimate for the included district--4.43 as against 5.93 children per women, respectively. The quinquennial age-specific birth rates for Kuala Langat in Table 9 are obtained by averaging the single year own-children estimates in Table 8. The own-children rates are probably somewhat low, if anything, on account of census underenumeration of children aged 0. The rates calculated from birth registration data are probably somewhat low also on account of under registration of births. Estimates for Peninsular Malaysia as a whole give birth registration completeness of over 95 percent, however, and it is scarcely plausible that under registration in the district of Kuala Langat could be so great as account for the difference between 4.43 and 5.93. We therefore conclude, at least provisionally, that the difference between the two figures reflects a difference in fertility between the state of Selangor and its administrative district Kuala Langat. The implication is that the own-children estimate for Kuala Langat is approximately correct and that the much lower fertility in Selangor is due to much lower fertility in the rest of the state.

In view of this substantial fertility differential, one would like to have a direct estimate, based on birth registration data, for Kuala Langat, and some progress in this direction may be made indirectly by utilizing the data on births by district of occurrence. The relevant figures are shown in Table 10. The age-distribution of women aged 15-54 in Kuala Laugat is available from the own-children tabulations (column 1). If age-specific fertility rates in the district were the same as those in Selangor State (column 2)--a supposition we have provisionally concluded to be false--there would have been approximately 2,910 births during 1970. Since age-specific fertility rates in the district are, we think, higher than those of Selangor, registered births in the district ought to exceed

this figure. The actual number registered accordingly comes as a disconcerting surprise—2,740 births!

At first glance, this leaves us with two obvious alternatives, an own-children estimate on the order of 30 percent too high, or birth under-registration of about 30 percent, both for Kuala Langat district. Neither alternative is plausible, for the reasons referred to above. There is of course the distinction between classification by district of occurrence of birth and district of residence of mother. Is it possible that a substantial number of births to women living in Kuala Langat occurred outside this district? This might be the case if a considerable proportion (at least 30 percent) of women give birth in hospitals, and if there are no suitable hospitals in the district. Indeed, the two previous alternatives are so unpalatable that one might go so far as to predict the absence of hospital facilities in Kuala Langat. In any case, it happens that there is in fact no district hospital in Kuala Langat (a fact for which I am indebted to Rabieyah Othman Mat, Statistician for Social Statistics, Census and Demography Division, Department of Statistics, Malaysia). As for confinements in hospitals, data on hospital admissions in Peninsular Malaysia for 1969, 1970, and 1971 show maternity admissions in the numbers of 121,884, 116,186, and 124,239, respectively (Social Statistics Bulletin: Peninsular Malaysia 1969–1971, Table 6, pp. 170–172, Department of Statistics, Kuala Lumpur, Malaysia). Registered births during these years under 297,693, 297,358, and 309,378, respectively (Vital. Statistics: Peninsular Malaysia 1972, Table 14.01, p. 33, Department of Statistics, Kuala Lumpur, Malaysia). Births occurring in hospitals may accordingly be estimated as 40 percent, 39 percent, and 40 percent respectively. We conclude that the relatively low incidence of births occurring in Kuala Langat is accounted for by women traveling outside the district to give birth in hospitals in other districts.

12. ACKNOWLEDGEMENTS

I am grateful to R. Chander, JMN, Chief Statistician, Malaysia, to Dorothy Z. Fernandez, Senior Statistician, Census and Demography Division, Department of Statistics, Malaysia, and Dr. Eduardo Arriaga, International Statistical Programs Center, U.S. Bureau of the Census, for the opportunity to present this material to the workshop; to Dorothy Fernandez, additionally, for providing the special census tabulation on which the example is based; to Shaari Abdul Rahman, Statistician for Evaluation, Vasantha Kumary Kandiah, Statistician for Fertility, and

Rabieyah Othman Mat, Statistician for Social Statistics, all of the Census and Demography Division, for numerous helpful conversations; and to all the participants in the workshop for their questions and comments, which have resulted in several improvements of these notes over the original lecture. I am further grateful to Lee-Jay Cho, Director of the East-West Population Institute, for providing information concerning the historical development of the method.