PARITY PROGRESSION PROJECTION

Griffith FEENEY
East West Population Institute,
Honolulu, Hawaii

In a previous paper (Feeney, 1983) I have shown how population dynamics may be formulated in terms of parity progression. The present paper develops techniques for carrying out population projections based on parity progression dynamics.

PROJECTING FIRST BIRTHS

To project first births we apply conventional component projection to the population of zero parity women. We begin with a single year age distribution of zero parity women at some time t=0 and a schedule of first birth rates, defined as zero parity women aged x at the beginning of any year divided into the number of these women who have a first birth during the year. Multiplying the initial number of zero parity women aged x by the corresponding rate thus gives the number of these women who have a first birth during the year, and subtracting this number from the original number gives the number of zero parity women aged x+1 at time t= 1 . Doing this for all ages, we obtain a projected age distribution of zero parity women for time t=1 and (summing over ages) projected first births for the time period (0,1). Mortality, ignored here, may be handled in the usual way.

TRENDS IN PROGRESSION TO FIRST BIRTH

For the United States, first birth rates ("probabilities") with zero parity women in the denominator may be obtained from the first three pages of Table 10A of Heuser (1976). The entry of this table for each time t (January 1 of each year) and age x gives the proportion of zero parity women aged x-1/2 to x+1/2 at time t who have a first birth before time t+l. Values for women aged x in completed years may be obtained by averaging. For some purposes it is more convenient to work with the five year age group values given in Table 9A. The entry for each time t and age group x to x+n in this table is the proportion of zero parity women aged x-1/2 to x+n-1/2 at time t who have a birth during the following year.

A period measure of the proportion of women ever progressing to a first birth may be defined by analogy to period life table statistics. We ask, for a hypothetical birth cohort of women experiencing the first birth probabilities observed in a given year, what proportion would eventually progress to a first birth. This proportion is given by one minus the product over all ages x of one minus the first birth rate for age x . Given first birth probabilities for five year age groups, one may simply impute the probability for the group to each single year of the group and proceed as with single year values.

PROJECTION WITH CHANGING SCHEDULES

To conveniently specify a series of changing schedules we convert the schedule of first birth probabilities into a schedule of forces of progression to first birth, defined analogously to the force of mortality, and assume a "proportional hazards" model, multiplying the force schedule by a constant to raise or lower the level of progression to first birth. Assuming the force of progression to be constant in the age interval x to x+1 , the probability p_x that a zero parity woman aged x exactly will progress to a first birth within one year may be expressed as 1-exp(-n_x) , where n_x, denotes the force of progression to first birth, so that n_x = -ln(l-p_x) . Because 1-z approximates exp(z) for small z, the force and the probability are approximately the same if the probability is low.

Beginning with a base schedule of probabilities for progression to first birth, then, we calculate the corresponding schedule of forces. To obtain first birth probabilities for any given future year, we multiply the force schedule by a certain constant k and convert back to a schedule of birth probabilities. The natural way to specify k is to specify a period parity progression ration p_0 for progression to first birth and then find the value of k that gives this p_0 Since p_0 = 1 - exp{k? n_x}, k = {-ln(1-p_0)}/{ ? n_x } . Under this model, schedules of progression to first birth for a series of years are determined by a single "standard" schedule and a series of parity progression ratios for progression to first birth.

INDIRECT ESTIMATION OF PROGRESSION TO FIRST BIRTH

Suppose we have a single year schedule of first birth probabilities for 1973 and an age distribution of zero parity women as of January 1, 1974, but that we also have total numbers of first births for more recent years, through 1981, say. This is, in fact, precisely what we do have for the United States at this writing, a result of different delays in the availability of different data.

Imagine that we project forward one year, obtaining a projected number of first births for 1974 and an age distribution of zero parity women as of January 1, 1975. The number of first births we project may be raised or lowered by raising or lowering the schedule of first birth probabilities, which ore determined by the value of p_0. We may therefore obtain any number of projected first births we wish by selecting the right value. In particular, we may select the value of p_0 that gives the observed number of first births, and in this way we obtain an indirect estimate of the schedule of birth probabilities for 1974. The same process may then be repeated for each succeeding year for which we have an observed number of first births.

This estimation procedure is formally analogous to what Lee (1974) has called "inverse projection". It will be used again below in connection with progression to higher order births.

PROGRESSION TO HIGHER ORDER BIRTHS

We put projection proper aside for a moment to consider the measurement of progression to second and higher order births. The essential ideas go back to Henry (1953: chapters 2-3).

The topmost panel of table 1 shows data on progression to second birth, for parity cohorts (Feeney 1983: introduction and the section beginning on page 83), in Lexis diagram format. The first row of this panel gives the numbers of first births by year of occurrence. These first births represent entries, analogous to "births," to the population of parity one women. The following rows give numbers of second births to these women In successive years. These second births are classified by year of occurrence and parity status of mother, with “Entries” indicating second births to women who had a first birth in the same year, and “CYDIP = i’ indicating second births to women with i completed years duration in parity at the beginning of the year. These second births represent departures from, analogous to “deaths to,” the population of parity one women.

Regarded from the cohort perspective, the data indicate that, for example, of the 1122 women who had a first birth in 1960, 12 had a second birth in the same year, 322 had a second birth in the following year, and so on. Regarded from the period perspective, they show that, for example, of all second births occurring in 1965, 13 occurred to women who had their first birth in the some year, 267 occurred to women who had their first birth In the preceding year, and so on.

The second panel of the table shows open birth interval distributions (see Feeney and Ross, 1984, for a general discussion). These are derived as follows from the numbers of births given In the data panel. Of the 1159 women who had a first birth in 1965, 13 had a second birth in the same year, leaving 1146 parity one women with zero completed years duration in parity at the beginning of 1966. Of the 1167 women who had a first birth in 1964, 14 had a first birth In the same year, leaving 1153 parity one women with zero completed years in parity at the beginning of 1965, and another 267 had a first birth in 1965, leaving 886 parity one women with one completed year duration In parity at the beginning of 1966.

The third panel shows rates of progression to second birth for parity one women. The values in the first row are obtained by dividing the number of second births each year to women who have a first birth In the same year by the total number of first births in this year. For 1965, for example, we have 13/1159 = .011 .The values in subsequent rows are numbers of second births each year to parity one women with a given number of completed years duration in parity at the beginning of the year, taken from the data panel, divided by the number of these women, taken from the open birth Interval distribution panel. Thus for example the rate for women with two completed years duration in parity one at the beginning of 1965 is 173/520 = .333.

The fourth panel shows period parity progression ratios. As with progression to first birth, these are defined by analogy with period life table statistics. Consider an hypothetical cohort of 1000 women who have a first birth in a given year, and suppose they experience the rates of progression to second birth shown in the “Rates of Progression” panel for 1965. We see at once that the proportion of these women who will hove progressed to a second birth by the end of the given year will be .011, that the proportion who will have progressed to a second birth by the end of the following year will be I -(1 - .011)(1 - . 232) = .240, and so on.

The proportion of women who progress to a second birth within five years duration in parity, the “quinum” of Rodriguez and Hobcraft (1980), may be approximated by averaging the proportions of women who progress to a second birth by the end of the fourth and fifth year following the year of their first birth. The value for 1965 would thus be (0.773 + 0.805)/2 = 0.789.

As the values in table 1 are merely illustrative, only a few values ore given. In practice one will usually want data for as many years as one can get, and one will want to carry the duration in parity values out to ten or more years.

Smoothing may be effected with minimal sacrifice of time trend detail by "scrolling," carrying out calculations for moving aggregates of two or more years. To illustrate, the aggregate rate at which women who had a first birth in 1964-65 had a second birth in the same year is, referring to the numbers In the data panel of table 1, (14 + 13)/(1167 + 1159) = 0.012. The rate at which women with zero completed years in parity at the beginning of these years proceeded to a second birth is (276 + 267)/(1115 + 1153) = 0.239, and so on. The same idea may be carried through all the calculations in the table. This is useful when dealing with survey data, where small numbers give rise to random fluctuations in numbers of births.

PROJECTING HIGHER ORDER BIRTHS

To project second births we again apply conventional component projection, this time to the population of parity one women, and similarly for higher birth orders. The initial "age" distribution is an open birth interval distribution, and projection consists of applying rates of progression, as shown in table 1, to this distribution.

The process is readily illustrated by example. Suppose that, given the data in table 1, we wish to project second births during 1966, assuming that rates of progression in this year will be the same as those observed in 1965. Second births in 1966 contributed by women who were parity one at the beginning of the year are calculated as .232 times 1146, plus .330 times 886, and so on. Second births in 1966 contributed by women who entered parity one status during the year are calculated as the number of these women, which equals the number of first births during the year, times the corresponding rate, 0.011. Summing these two components gives total projected second births for the year.

As with progression to first birth, it is simple enough in principle to change the schedules of rates from one year to the next, but in practice we want a simple way of specifying a set of changing schedules. The proportional hazards approach described above may be applied equally well here. A modification might be mode to account for the incomplete exposure attached to successive births occurring in the same year, but the magnitudes involved will usually be considered too small to justify the trouble. Alternatively, we might simply multiply a single "standard" set of rates by a constant factor to change the rates from one year to the next. We shall use this approach in the indirect estimation described below.

OVERVIEW

The input data required for a single cycle of parity progression projection are (1) the single year age distribution of zero parity women under age 50 as of time zero, (2) a corresponding single year schedule of first birth probabilities, (3) open birth interval distributions for parities one two, ..., and (4) corresponding schedules of rates of progression to next birth. The age and open birth interval distributions describe the state of the population at a particular point In time, and the corresponding schedules of rates stipulate how this state will change over the following year.

The results of a single projection cycle are (a) projected births by order and (b) updated values of the age and open birth interval distributions, i.e., values that describe the state of the population at the end of the year. If the projection is made in conjunction with an ordinary component projection, one may also obtain (c) projected age-specific first birth rates and (d) projected proportions of zero parity women by age.

Projection with changing rates requires the distributions (1) and (3) for the initial "time zero" and the schedules of rates (2) and (4) for every following year. Invoking the "proportional hazards" model, the specification of these schedules of rates may be simplified to (i) a single set of "standard" schedules and (ii) a set of parity progression ratios for each year for which the projection is to be carried out. Projection with constant rates requires only items (1-4). The results of projection, in each case, are time series of the values (a-b), or (a-d), of the last paragraph.

SOURCES OF DATA

The ideal source would be a birth registration system that collects birth order for all live births and year of last birth for second and higher order births. One would then tabulate all live births each year by birth order and second and higher order births by birth order and year of last birth. Registration must be complete, however, and we must have a reasonably long series of data, ten or more years, say.

An alternative and more widely available source is a fertility survey including a birth history. To obtain data on progression from first to second birth one first tabulates all women with one or more children by year of first birth. This gives numbers of survey women having a first birth each year, corresponding to the first row of the data panel of table 1. One then tabulates all women with two or more children by year of first birth and year of second birth minus year of first birth. This gives the numbers of women having a first birth in each year who progress to a second birth in each subsequent year, corresponding to the lower rows of the data panel in table 1. Similar tabulations may be made for higher order progressions.

When two or more fertility surveys are taken a number of years apart, the retrospective estimates from the latter surveys will overlap the retrospective estimates from the earlier surveys. If the estimates are correct, the overlapping portions will be Identically equal. Comparison of overlapping trends thus provides a powerful consistency test with which to check data quality.

A disadvantage of survey data on birth histories is the age-selection bias Introduced by mortality and the usual restriction to reproductive age women. A survey of women under age 50 taken in 1965 provides some information on progression to second birth for women who had their first birth in (say) 1945, but it obviously excludes women who had reached age 50 or died before the survey was taken. The survey women who had a first birth in 1945 were thus substantially younger on the average than the group of all women who had a first birth in this year, and it is reasonable to expect that this selection will bias any parity progressions based on such data (Rindfuss, Paimore and Bumpass, 1982).

A second disadvantage of birth history surveys is that they are often rather small for the uses described here. If registration data on births by order is available, both of these disadvantages may be largely overcome by applying the method described in the following section.

SIMULTANEOUS ESTIMATION FROM SURVEY AND VITAL REGISTRATION DATA

Suppose we have available (a) a long series of birth registration data giving annual births by birth order and (b) a birth history fertility survey. We may make a trial estimate of the number of second births in year y+t to women who had their first birth in year y by multiplying the registered number of first births in year y by the proportion of survey women having a first birth in year y who had a second birth in year y+t . Carrying this out for a series of years yields a table of trial estimates of second births like that shown in the data panel of table 1. Summing down columns in this table for any given year gives a trial estimate of second births during this year. These trial estimates may be reconciled with the vital registration data by multiplying the numbers in each column by the ratio of registered second births to the sum of trial estimates.

If both sources were free of error, these consistency ratios would all be one. Random fluctuations due to small numbers in the survey will cause irregular fluctuations above and below one. If the survey numbers are significantly affected by random variation, it may be desirable to scroll the data over two or more years calculating and applying the consistency ratios.

Since younger women will in general show higher progression to second birth than older women, we would expect age-selection bias to yield over-estimates of second births from the survey data, the more so as we move further back from the time of the survey. By forcing the survey data into conformity with the vital registration totals, this procedure controls for age-selection bias. A declining trend in the consistency ratios as we move back from the survey may be interpreted as a measure of the quantitative effect of age selection bias.

INDIRECT ESTIMATION OF PROGRESSION TO HIGHER ORDER BIRTHS

Suppose we have available a long series of annual numbers of births distributed by order, but no direct data on parity progression. For specificity consider progression from first to second birth. The available data thus consists a long series of annual numbers of first and second births beginning in some year y. Suppose for the moment that we are able to obtain two additional items, an open birth interval distribution for parity one women at the beginning of year y , and a set of “standard” rates for progression of parity one women to second birth. Multiplying the standard rates by a suitably chosen factor will yield a set of rates which, when applied to the initial open birth interval distribution, will give the observed number of second births. This same set of rates may then be applied to compute the open birth interval distribution for parity one women at the beginning of year y+1 . The same process may then be repeated for year y+1, and for every succeeding year to the end of the available series.

In this way we reconstruct a complete set of parity progression data for the entire length of the series. The extent to which these numbers estimate the true population numbers depends on two factors, the accuracy of the open birth Interval distribution we begin with, and the extent to which actual rates of progression in the population conform to the simple “multiplication by a constant” model assumed in the calculation. This is in addition, of course, to the accuracy of the given birth order data.

Given the ergodic theorems of demography (see for example, Arthur 1982), It is natural to ask whether the initial open birth interval distribution is "forgotten" as the process proceeds. Trial calculations suggest that, in this application at any rate, forgetting not only occurs, but occurs very rapidly. Beyond the first few years of the process, which can be continued for some 50 years with data available for the United States, the estimated parity progression values are extremely robust against variations in the initial open birth interval distribution. The rapidity of the forgetting no doubt reflects the short average interval between births. Wachter (1984) provides interesting and important results on ergodic theorems in a similar context.

The dependence of the results on the rates model assumed is more problematic. The distribution of the reconstructed data by duration in parity obviously depends strongly on the pattern of the standard schedule of rates. Although we will presumably be able to choose a sensible pattern even in the absence of any direct data on the population in question, it would be unwise to lean heavily on the duration detail in the reconstructed data.

In any case, our principal interest is likely to be the overall level of progression from first to second birth, as indicated by, say, the period parity progression ratio at ten years duration in parity. We would not be surprised to find these values relatively robust against variations in the standard schedule of rates, but detailed results must await detailed numerical computation, as well as accumulation of evidence on empirical patterns of rates of progression.

Even if derived parity progression ratios turn out to be highly robust against variations In the standard, however, there is the possibility that the changes in the rates from one year to the next are not well represented by multiplication of a standard schedule by a constant. Initially at least, the best way to explore this issue Is probably to study populations where the indirect estimates may be compared with direct estimates from survey data. This is possible for any population for which we have available a reasonably long series of data on annual births by order and one or more birth history fertility surveys.

The formal demographic structure of the situation suggests that this procedure will give better results that the indirect procedure proposed by Henry (1953, chapter 3), but it would be instructive to pursue the comparison.

CONCLUSION

Population projection has been formulated in terms of period parity progression ratios, as opposed to the usual formulation in terms of period age-specific birth rates. In analytical applications, this allows us to phrase and answer new questions about fertility trends and population growth. We may for example assess the effect of trends in progression to first birth on fertility trends by asking what the trend of fertility would have been if progression to higher order births had remained constant while progression to first birth followed the course that it did follow (cf. Ryder, 1980). The birth planning efforts In China provide another example of an area for analytical applications.

In forecasting applications, parity progression measures of fertility provide an alternative way of locking at past fertility experience, and hence on alternative way of extrapolating post experience into the future. The difficulties of forecasting are of course due in considerable part to the uncertainty inherent in the phenomena, a fact that no innovation in technique is likely to overcome. Hajnal's (1955) remarks on forecasting are as pertinent today as when they were written 30 years ago, and his pessimism is by and large supported by the recent work of Brass (1974) and Keyfitz (1981, 1982). Past forecasts have been so far wide of the mark, however, that even major improvements would be consistent with a continuing substantial element of uncertainty.

An essential difficulty of assessing fertility trends consists in the impossibility decomposing current changes in fertility into components that reflect (a) changes in the timing of births, which are necessarily transient and reversible, and (b) changes in the level of completed fertility, which are potentially persistent. Our recognition of this distinction is due to Ryder (see for example 1980 and the references therein). Current data can tell us only that, say, women currently reaching age 25 have had fewer children than women who reached age 25 In the past. Twenty-five years from now, when these women hove reached the end of reproduction, we will be able to look back and compare the level and timing of their fertility with that of previous cohorts. For the present, however, any judgment as to how this comparison will turn out is necessarily a prediction of what these women will do in the future.

Changes in the timing of childbearing consist of changes in the interval to first birth and of changes in interbirth intervals. Changes in the lnterbirth intervals are relatively unproblematic simply because the intervals are short relative to the time spans with which we are concerned. Shifts in the timing of progression to next births can't mislead us for more than a few years, because nearly all women who ever progress do so fairly rapidly. By contrast, the interval between the birth of a woman and the birth of her first child is so long that gradual shifts in the timing of first birth may create spurious fertility trends that persist for many years.

Decomposing overall fertility into progression to first birth and progression to higher order births effectively isolates the timing versus completed level dilemma to trends in progression to first birth. This frees us to analyze trends in progression to higher order births relatively unencumbered. It may also make the dilemma easier to cope with, for the social and economic considerations that influence progression to first birth are evidently narrower than those that influence fertility at large. Finally, progression to first birth based on birth "probabilities" is relatively insensitive to shifts in the timing of first birth. Lower rates of progression at young ages, where we expect timing changes to have most effect, will result in larger numbers of zero parity women at higher ages. Since our rates are applied to zero parity women, this will result in larger numbers of first births at older ages, partially canceling out the decline at younger ages.

SUMMARY

A method of population projection based on parity progression is developed. First births are projected by applying suitably defined age-specific first birth rates to an initial age distribution of zero parity women, second births by applying duration-in-parity-specific second birth rates to an initial open birth interval distribution of parity one women, and so on for third and higher order births.

A proportional hazards model for projection with changing schedules of progression to first birth and an indirect procedure for estimating these schedules are proposed.

Procedures for obtaining parity progression data retrospectively from fertility survey birth history data are given. When two or more surveys are available, comparison of the overlapping portions of retrospectively estimated trends provides a powerful check on data quality.

A method for estimating parity progression data simultaneously from fertility survey birth, history data and vital registration data on births by order is presented. This provides another way of checking data quality, reconciles the two sources of data, and alleviates the small number and age selection bias problems present in most survey data.

When registered births by order are the only source of data, but are available for a reasonably long period, parity progression data may be estimated indirectly by Inverse projection.

The conclusion identifies applications and potential advantages of parity progression projection In analytical and forecasting applications.

REFERENCES

Arthur, W.B., 1982. The ergodic theorems of demography: a simple proof, *Demography*, Vol. 19, No. 4.

Brass, W., 1974. Perspectives in population prediction: illustrated by the statistics of England and Wales, *Journal of the Royal Statistical Society*, Series A, Vol. 137, Part 4.

Feeney, G., 1983. Population dynamics based on birth intervals and parity progression, *Population Studies*, Vol. 37, No. 1.

Feeney, G. and Ross, J. .A., 1984. Analyzing open birth interval distributions, *Population Studies*, forthcoming, Vol. 38, No. 3.

Henry, L., 1953. *Fertility of marriage: A new method of measurement*, Population Studies Translation Series No. 3, United Nations, 1980 (French edition published 1953).

Heuser, R. L., 1976. *Fertility Tables for Birth Cohorts by Color: United States, 1917-73*, National Center for Health Statistics, Rockville.

Keyfitz, N., 1981. The limits of population forecasting, *Population and Development Review*, Vol. 7, No. 4.

Keyfitz, N., 1982. Can knowledge improve forecasts?, *Population and Development Review*, Vol. 8, No. 4.

Lee, R.D., 1974. Estimating series of vital rates and age structures from baptisms and burials: a new technique, with applications to pre-Industrial England, *Population Studies*, Vol. 28, Part 3.

Rindfuss, R.R., Palmore, J.A. and Bumpass, L.L., 1982. Selectivity and the analysis of birth intervals from survey data, *Asian and Pacific Census Forum*, Vol. 8, No. 3.

Rodriguez, G., and Hobcraft, J., 1980. Illustrative analysis: life table analysis of birth intervals in Colombia, World Fertility Survey, London.

Ryder, N.B., 1980. Components of temporal variations in American fertility, in R.W. Hiorns, ed., *Demographic Patterns in Developed Societies*, Taylor & Francis, Ltd., London.

Wachter, K. W., 1984. Ergodicity and inverse projection, typescript, Demography and Statistics, University of California, Berkeley.