NOTES ON FERTILITY MEASUREMENT

NORMAN B. RYDER

The realm of fertility measurement has so considerably broadened and deepened during the past forty years that any survey bounded within a reasonable number of pages becomes either a superficial listing or a nonrandom selection. The present paper follows the latter course. Although an attempt has been made to exemplify characteristic problems and solutions in the diverse parts of the subject, the treatment is evidently parochial—confined to work in and data about the United States, and biased in the direction of what the writer considers to be important past and present issues. The magnitude of the assignment is partial justification for this implicit presumption.

It is convenient to distinguish two polar types of analysis toward either of which the task of fertility measurement may be primarily oriented. On the one hand is causal analysis, consisting of attempts to provide explanations for fertility behavior as the dependent variable. Most of the data for this purpose have until recently come from the official systems of registration and enumeration; this is the topic of the first section below. These can now be supplemented considerably by detailed reproductive histories collected in surveys by nongovernmental researchers—because officials tend to be nervous in the presence of intimate secrets. Measurement problems associated with this means of obtaining data are discussed in the second section of the paper.

The other type of inquiry about fertility is directed to questions of consequential analysis. Here the demographic inputs play the role of independent variables in the models of economists and others concerned with the impact of population change on the variables that define their disciplines. Although the characteristic style of causal analysis is decomposition of individual processes, from a longitudinal perspective, the consequential format tends to be aggregative and cross-sectional, and primarily oriented to variations in total population structures. This is the subject of the final section of the paper.

Differences in style at these analytic poles provide one justification for the interstitial sector of formal demography. Inquiry within the formal sector does not share the substantive aspect of causal or consequential analysis; it consists of the purely deductive development of relations between measures designed for one purpose and those designed for another purpose. Whether the formal procedure is deterministic or stochastic, and whether it is accomplished by elegant derivation or pedestrian simulation, the intent and the result are mathematical. Yet this is far from implying practical irrelevance to such work. Policy-making requires causal analysis, to locate the crucial and modifiable facets of the processes that generate possible futures, and consequential analysis, to identify among them the future that is the least undesirable; formal demography provides the dictionary for translating both ways between causal inputs and consequential outputs.

CAUSAL ANALYSIS BASED ON OFFICIAL STATISTICS

Most of the data employed in demographic analysis are perforce the by-products of procedures developed primarily with consumers other than demographers in mind; viz., the official systems of registration and enumeration. By the nature of the case, these data come to us in consequential form, and much labor is needed to transform them into fit shape for causal inquiry. The prototypical measure is the age-specific birth rate. From a causal standpoint, it represented a first step in the direction of precise identification of exposure to risk; from a consequential standpoint it was the linchpin of the stable population model, discussed below. During the 1930's and 1940's, the capabilities of this approach were developed in two directions: generalization of the principle of specification; and recognition of the relevance of the distinction between alternative modes of temporal aggregation of vital rates.[1] Where official systems made this feasible, birth rates were calculated specific not only for age but also for parity, for marital duration and for interval since preceding birth. It is technically feasible to characterize a set of birth rates specific for all such variables simultaneously, derive the exposure distributions that would eventually ensue were such rates to persist indefinitely and employ such distributions as weighting systems for the rates, preferable in general to the weighting systems implicit in the population being observed. For each particular set of specificities there is a distinctive reproductivity process (and thus a distinctive stable population). In retrospect this elegant formal accomplishment, to which Wicksell, Quensel, Vincent, Hyranius and Whelpton, as well as the redoubtable Lotka, contributed, has proven to play a negligible role in causal analysis of fertility—although it may have helped to disseminate the fruitful idea of a population in a stable state among social scientists whose data could be put into renewal form.

The second direction of development was a growing understanding of the significance of the two time vectors of fertility measurement—the period and the cohort. Demographic data are naturally produced period by period, and consequential analysis ordinarily uses a model with period-specific demographic inputs. The employment of period-specific indices for causal analysis, on the contrary, has foundered on the elementary but fundamental circumstance that the individuals whose behavior is to be explained in fact age year by year—they follow the cohort vector. It is tempting to propose that cohorts are the natural units of causal analysis and periods the natural units of consequential analysis, with the corollary that one essential facet of formal analysis is the derivation of formulae for translation between cohort and period modes of measurement. That this is a simplistic generalization is evident on the one hand in consideration of the many events such as legislation (the recent revocation of the liberal Rumanian abortion code for example) or technologic development (such as the production of the oral contraceptive), which impose a distinctive period-specific mark on the fertility surface, irrespective of cohort identity. It is evident on the other hand in recognition of the continuity of consequences of demographic change, along the cohort diagonals, despite an initial periodic configuration.

The major figure in the development of cohort analysis, although there were important predecessors, was P. K. Whelpton.[2] It is an interesting sidelight on this history that Whelpton first organized his data in cohort form for the purpose of obtaining the exposure denominators needed to calculate age-parity-specific birth rates, and only subsequently perceived that the rearrangement he had adopted essentially for accounting purposes was in fact a much more fundamental contribution to fertility measurement than the increased specificity that was his initial target.

The cohort fertility tables for the United States, for each year since

TABLE I. PARAMETERS OF COHORT FERTILITY, UNITED STATES

Birth Cohort Group	*Total Fertility Rate*	*Mean Progression Ratio*		*Mean Age at Marriage*	*Mean Birth Interval*
		Lower Parity	*Higher Parity*		
1896–00	2.675	0.866	0.662	20.89	2.57
1901–05	2.421	0.855	0.630	20.63	2.84
1906–10	2.273	0.849	0.606	21.23	2.94
1911–15	2.316	0.863	0.591	21.76	2.97
1916–20	2.553	0.893	0.594	21.37	3.05
1921–25	2.864	0.922	0.611	20.75	3.01
1926–30	3.092	0.932	0.635	20.43	2.64
1931–35	3.293	0.947	0.637	19.96	2.41
1936–40	3.079	0.946	0.612	19.42	2.57
1941–45	2.779	0.920		19.58	

1917, provide birth rates specific for age and birth order, for each single-year cohort.[3] This sophisticated set of information represents a substantial advance, both in idea and in execution, over previous compilations. Its proper use for contemporary analysis, however, faces one obstacle that is in a general sense intractable. Conventional summary indices, such as the moment measures of the fertility-age function, can be used routinely to characterize the experience of any cohort that has already completed childbearing, but the cohorts that are of highest current interest are still in reproductive transit. One cohort may have less fertility currently than a predecessor at the same age either because its eventual total will be smaller, or because it has so far borne a lower proportion of a larger eventual total.

Elsewhere the writer has proposed a procedure for making undistorted estimates of parameters of completed fertility, for truncated histories.[4] The validity of the procedure depends on the assumption that cohorts share a common age-pattern of fertility, provided age is transformed into units of standard deviation above or below the mean. This is a tolerable assumption for available American experience, so long as the point of truncation is at least as high as the putative eventual mean. This procedure has been used in making the calculations for recent cohorts in Table 1.

A second difficulty may be remedied with the right kind of information, but that is still unavailable. There is no beginning to the reproductive sequence displayed in the cohort fertility tables, because there are no usable nuptiality data. Furthermore, it is not possible to ascertain how long a time elapses between one birth and the next without specification of this information on the registration form. The difference between mean age at Nth birth and mean age at N+1th birth is of little help, because it is likely that a difference exists in mean age at Nth birth between those who do advance from Nth to N+1th birth and those who do not. Some assumptions about this difference are necessary to estimate mean age at first marriage and mean length of birth interval from the cohort fertility tables as presently constituted.

One model that accomplishes this is the following:[5] It is advantageous to study order-specific total fertility rates by converting them into parity progression ratios; i.e., the probabilities of having an N+1th birth, conditional on the occurrence of an Nth birth. The proper beginning for the sequence is the probability of first marriage (prior to the end of the childbearing period), say P(0), followed by the probability of a first birth, given that marriage has occurred, P(1), and so forth. The evidence for the United States is that the time series for P(N), where N LE 2, have followed one pattern, and that the time series, where N GE 3, have followed another pattern. Given this circumstance, as well as the salience of the two-child family (the point of dichotomization of the two patterns) both for reproductive norms and for the level of fertility requisite to replacement, a model is suggested, for purposes of analytic summary, in which the lower-parity progression ratio has the value L, and the higher-parity progression ratio the value H. If the total fertility rate observed for the first two orders of birth combined is F(1, 2), and that for the higher orders of birth is F(3+), then an estimate of L can be found by solving the equation $L^2 + L^3 = F(1, 2)$ and an estimate of H by solving the equation $H = F(3+)/(L^3 + F(3+))$. These are the values shown in Table 1 for the available American data.

Of more interest, because the information is more deeply hidden within the basic tables, are parameters of the time distribution of cohort fertility. Assume that the unknown mean age at first marriage is A(0). Marriages are divisible into those that do and those that do not produce a first birth, identified symbolically by a single and a double prime respectively. $A(0) = (L)*A'(0) + (1-L)*A''(0)$, since in the model, L is the proportion progressing from marriage to first parity. The tables permit calculation of the mean age at first birth, A(1); it too is divisible into two components—a mean age for those who go on to have a second, A′(1), and a mean age for those who do not, A″(1), so that $A(1) = (L)*A'(1) + (1-L)*A''(1)$. The same holds for each higher order of birth. At each successive parity, there are two variables of note: the length of the birth interval, for example $I(1, 2) = A(2) - A'(1)$; and the difference between the mean age for those who do and the mean age for those who do not progress to the next parity, for example $J(2) = A''(2) - A'(2)$. With the intention of achieving average values, let $I(N, N+1) = I$ and $J(N) = J$; to make the entire system determinate (without gross departure from fact), let $I = J$.

With this assumption, the mean age at first marriage, $A(0)$, and the mean birth interval, I, can be estimated as follows: Let the mean age of fertility of the first and second order be $M(1, 2)$ and the mean age of higher order fertility be $M(3+)$. Then $I = (M(3+) - M(1, 2))/(C-B)$, and $A(0) = (C^*M(1, 2) - B^*M(3+))/(C-B)$ where $B = 2L - L/(1+L)$ and $C = 2L + H/(1-H)$. The results are shown in the last two columns of Table 1.

The indices shown in Table 1 provide a compact summary of the 15,000 or so rates that constitute the cohort fertility tables for the United States. To interpret the implications for the total fertility rate of variations in the two mean progression ratios, it is helpful to know that the total fertility rate is equal to $L^2 * (1+L-H)/(1-H)$, and accordingly that, over the range of this series, the partial derivative of the total fertility rate with respect to L is almost twice that with respect to H. The principal contributor to the baby boom (when the total fertility rate rose from 2.273 to 3.293) has been L, the lower-parity progression ratio. However, the future will not show a repetition of this kind of baby boom (unless L first declines to its depression level) because L is a conditional probability that is evidently very close to its physiologic maximum. In this particular slice of American vital history, the higher-parity progression ratio, H, has shown rather little variation comparatively, although its current value is distinctly lower than at the beginning of the series. With a longer perspective, however, it would undoubtedly be the case that the secular decline in overall fertility was essentially a decline in H (from a value that must have been close to 0.9 at the beginning of the nineteenth century).

Timing variations, as captured in the indices $A(0)$ and I, have been substantial throughout the past five decades. Although the range of variation in the surrogate age at first marriage, $A(0)$, has been greater than the range for the mean birth interval, I, the latter predominates as a determinant of the time pattern of fertility because its variation is weighted by the number of birth intervals; i.e., by the total fertility rate. Timing variations are of interest not only in their own right as problematic facets of cohort behavior but also because they are the source of distortions of period indices relative to cohort indices. The overt baby boom; i.e., that manifest in such period indices as the crude birth rate, was as much a consequence of a shift from rising to falling time patterns as it was a reflection of an increasing quantity of cohort fertility. The current period is characterized by decline in both of the quantity indices and rise in both of the time pattern indices; each has contributed approximately equally to the unprecedented collapse of the American birth rate during the 1960's.

Contrivances like the above model would of course be unnecessary if the requisite information about the timing of marriages and births were part of the registration system. Indeed a close approximation has been achieved by the Bureau of the Census, particularly through the device of the Current Population Survey. It would seem presently feasible and desirable to redo the cohort fertility tables, not only coordinating the data from registration and enumeration sources but also expanding the referent of the tables from total females to whites and nonwhites separately (and perhaps other characteristics that are suitable for cohort identification because they are fixed at birth, such as native-born and foreign-born).

One further modification may be recommended, not for its empirical impact—which is undoubtedly minor—but in the interests of methodologic purity. The basic principle underlying a preference for the cohort over the period mode of temporal aggregation is that parameters of the latter type are distorted versions of parameters of the former type whenever the age distribution of cohort childbearing is experiencing modification. Now each component of the cohort fertility tables; i.e., the rate (for each order) in each age, is as presently calculated a period surrogate that manifests its own miniature internal distortion. Any one of several simple interpolation procedures can remedy this defect. It may even be worth considering abandonment of age-specific calculations for cohorts in favor of period-specific calculations, on the ground that the major dimensions of variation in the period-cohort-age triad are the first two, with the diffuse effect of age quite adequately conveyed by an examination of the performance of a cohort from year to year throughout its childbearing span. Such a reorientation has obvious advantages in the solution of projection problems, as reported elsewhere.[5]

Whatever modifications may be made in cohort fertility tables, their usefulness is essentially limited to the most general appraisal of the character of temporal variations in American fertility—as an antidote to simplistic period-bound misinterpretations. By their very nature, they can provide little scope for the type of causal analysis aimed at establishing relations among the characteristics of individuals, because the available stock of information about each individual is so limited. To get around these constraints, there is considerable temptation to infer propositions about individuals from the free-floating analysis of the covariation of temporally coincident aggregate indices—essentially the temporal equivalent of ecologic correlations. The well-documented pitfalls of the approach have done little to deter work in this misguided tradition. A further difficulty, peculiar to cohort analysis of this sort, is that, inasmuch as the experience of each cohort covers a long span, no unique time point can be specified for any parameter that summarizes that experience.

No overriding scientific stricture dictates that the unit of analysis must be the individual (person or perhaps couple); the cohort can be treated as a meaningful analytic entity in its own right. Rich possibilities exist for research of this kind because essentially every survey specifies age as well as time of observation, and therefore inferentially cohort membership, but neither theorists nor researchers have so far taken up this challenge.

CAUSAL ANALYSIS BASED ON REPRODUCTIVE HISTORIES

During the past several decades the most important contributions to fertility measurement have come from those who have been developing models of reproductive behavior, especially Sheps and Perrin and Potter and Tietze. Their accomplishments are sufficiently numerous and complex to justify a paper devoted exclusively to that specialty; fortunately Sheps has recently provided such an essay.[6] The present account is restricted to exemplification of some distinctive aspects of this kind of work, using data from the 1965 National Fertility Study.[7]

In that survey (of a national sample of currently married women under the age of 55), we collected comprehensive pregnancy histories, including the use of contraception in each interval. Because the histories of the respondents were incomplete (truncated by interview) it was necessary, to obviate bias, to study experience in the open interval as well as in the closed intervals (those ending in a pregnancy) and develop month-by-month probabilities of "survival" (avoidance of pregnancy) to use in calculating indices. Table 2 shows some of these results.

The table contains results for the first five pregnancy intervals; i.e., for those intervals that ended in a pregnancy of the order indicated

TABLE 2. FERTILITY MEASURES FOR USERS AND NONUSERS, FIRST FIVE PREGNANCY INTERVALS, NATIONAL FERTILITY STUDY, 1965

Order of Pregnancy Interval	*Per Cent Using*	*Per Cent Surviving one year*	*Pregnancy Rate (%)*	*Per Cent Not Using*	*Per Cent Surviving one year*	*Pregnancy Rate (%)*	*Overall Pregnancy Rate (%)*
I	53	64	45	47	38	107	74
II	66	74	30	35	50	70	44
III	71	80	23	29	56	62	34
IV	70	80	22	30	57	55	32
V	68	78	24	32	60	51	33

(or would have done so, had the interval not remained open). The respondents are divided into users (at some time during the interval) and nonusers and, for each, two measures of fertility are presented—the proportion surviving without a pregnancy for at least one year, and the central pregnancy rate for the same experience. (These are the equivalents of what in a life table would be symbolized by ${}_1p_0$ and ${}_1m_0$.) The rate for users is a measure of contraceptive inefficacy; the rate for nonusers is a measure of fecundability. The difference between these is the justification for the specification of the status.

This measurement procedure differs in form and in substance from that used with official registration and enumeration data. The basic element of the latter is a central birth rate; i.e., a ratio of the number of births by persons within a specified category to the total number of person-years of exposure to risk within the same category. In Table 2, the basic event is pregnancy rather than birth, principally because the focus of the inquiry was the prevention of pregnancy by successful contraception (rather than the more inclusive topic of fertility regulation). A second difference is that the measure is interval-specific as well as order-specific; what is shown in Table 2 is only the first year in a sequence of rates or probabilities that can be calculated. As noted, interval-specificity is not feasible with the present registration data.

The final considerable difference is that it is feasible by means of reproductive histories to considerably sharpen the definition of exposure, by removing from the denominator of the fertility rate three categories of nonexposure: (1) the respondents who are sterile as a consequence of menopause, or sterilization of either spouse or morbidity; (2) respondents in intervals of pregnancy or puerperium; (3) those not exposed to risk because they were not married and living with husband (this serving as a substitute, neither necessary nor sufficient, for regular copulation). Parenthetically, this is a continuing difficulty for fertility surveys. The studies of 1955, 1960 and 1965 were restricted to women who at time of interview were married and living with their husband. This research design is justifiable on tactical grounds, but it excludes a lot of exposure. In the 1970 version the universe was enlarged to ever-married women. The problem remains of the proper measurement procedure for dealing with sporadic exposure (perhaps characteristic of nonmarital copulation), in contradistinction to that which has the continuous character suitable for a life-table approach.

Reproductive histories for cohorts that have not completed their childbearing pose methodologic problems similar to those noted previously for truncated cohorts in the cohort fertility tables, but there is more scope for their solution. If a cohort has completed its childbearing, the record of a reproductive history, in the form of a sequence of events of interest, by date, poses a straightforward problem of summarization, perhaps most conveniently handled by calculating the conditional probabilities of each event, given that the preceding event has occurred, as well as the time intervals between succeeding events for each respondent. For differentially incomplete cohorts, on the other hand, the best that can be done is to summarize the experience to date, in effect presenting the month-by-month probabilities of becoming pregnant (or remaining nonpregnant) for as long an interval as the data permit, and making comparisons (between cohorts, for example) only for standardized potential interval lengths.

One further precaution is necessary. In any cross-sectional survey by age, the representation of respondents in any interval is necessarily biased in favor of those respondents to whom the event that begins the interval occurs at an earlier rather than a later age. Inasmuch as the age at which each beginning event occurs, the "initial age," is ordinarily a strong correlate of behavior in the interval in question (as, for example, with age at first marriage), it is important in cohort comparisons to make the life table calculations separately for each initial age. The problem of estimating how two cohorts will compare in their eventual transition probabilities and interval lengths becomes twofold: (1) an estimate for each initial-age group of the eventual "survival" probability, essentially a problem of extrapolating the available curve; (2) an estimate for the cohort of its eventual distribution of each event by initial age (the weights for the estimates of eventual probability).

To summarize, the records of the times of occurrence of events of

TABLE 3. DISTRIBUTION OF RESPONDENTS IN INTERVAL III, BY INTENTION AND USE, AND PREGNANCY RATES (IN THE FIRST YEAR) FOR EACH CATEGORY

Intention	Distribution by Intent (%)	Users Per Cent	Users Pregnancy Rate	Nonusers Per Cent	Nonusers Pregnancy Rate
Nondelay	16			100	74
Delay	53	87	25	13	84
Termination	31	82	22	18	47

interest, like pregnancies, serve a function for the completed cohort of enrichment of the stock of information about the behavior. For the incomplete cohort they also provide the data needed for an essential control procedure, if bias is to be avoided. To return to a question raised in the preceding section, we would be in a substantially stronger position to make estimates of completed cohort fertility for cohorts currently within the reproductive age span if information were available about the interval since the preceding birth. The point being made is not at all peculiar to the study of fertility. It applies to the measurement of any event of status change, and therefore to the measurement of characteristics of persons occupying any status (except those fixed at birth) whenever the source of the data includes at least some respondents who will experience the status change subsequent to interview.

The usefulness of a reproductive history for causal analysis depends on the other information available for each respondent in the interval in question. Of particular theoretical interest is the intention of the respondent in each interval. Table 3 shows the results of the 1965 National Fertility Study for the interval between second and third pregnancy, with both use and intention specified.

In Table 3, those labeled "non-delay" are nonusers who gave as their reason for nonuse that they wanted a child as soon as possible. Those labeled "delay" indicated that they wanted a child eventually, but not as soon as possible. Those labeled "termination" said that they did not want another child at any time. In interpreting these data it is important to note that these intentions were retrospective reports; i.e., the respondent was reporting after the fact what the intention was before the fact. Among other obvious errors, this probably explains why the pregnancy rate for nonusers who said they intended termination was much lower than that for nonusers who had other intentions. It is doubtful that problems of this sort can be resolved without moving to a data-collection procedure that is longitudinal and repeated at short intervals.

A table with categories of intention and use implies a more clearcut identification of behavior than can be supported in actuality. The list of activities that affect pregnancy rates, but that may occur at least in part for other motives, is a long one. It includes "douche for cleanliness only," and its modern counterpart, "pill for medical reasons only," lactation and abstinence. Furthermore, given the prevalence of delayed marriage, sterilization (again frequently a question of mixed motives) and abortion, it is essentially unsupportable to identify the means appropriate to a particular intention as solely contraception. Finally it is not entirely clear what the referent of intention is. The idea implies a single actor, whereas it is quite clear that there are at least two actors in concert, and an array of other kin may play a role in the decision.

The fact is that the sequence, intention/action/outcome, implies a particular explanatory model that may or may not characterize most reproductive behavior in the United States. One hypothesis about reproductive behavior would identify the reproductive strategy in terms of a particular number of children toward which the appropriate actions are directed (and a particular time pattern for those children). This is the sense of a question as to how many children the respondent considers desirable. A second hypothesis, in the spirit of the questionnaires used in the 1965 and 1970 National Fertility Studies, is that the respondent has a meaningful judgment as to whether the next child is wanted, and if so, whether it is wanted right away or after some delay, but that a question concerning the total number wanted verges on meaninglessness because children are in fact born one at a time (with the trivial empirical exception of multiple births). This division of the number desired into a series of questions about whether the next is wanted is in the spirit of the parity progression ratio as a basic parameter of cohort fertility. The line of thought that carries us from desired family size to desire for the next child can be pushed one step further. It is not implausible that couples in fact decide only that they do or do not want a child to be the result of their sexual activities in a particular month, and that the question of whether they want another child eventually does not, need not, arise. It seems likely that many circumstances that have a bearing on the desirability of another child may change, and the labeling of any one interval by a single intention may blur this dynamic process. Certainly a reconsideration is appropriate if, within the interval in question, the respondent experiences a change of spouse. Finally the question may be raised as to the general appropriateness of a rational means-end schema for describing the behavior of all couples, even in a sophisticated modern society. This set of problems represents the frontier for causal analysis of reproductive histories, and little can now be asserted with confidence despite considerable accumulations of data.

MEASURES FOR CONSEQUENTIAL ANALYSIS

Progress in developing measures has been much less impressive for consequential than for causal analysis. Today, as forty years ago, the dominant indices are those associated with Lotka's model: the net reproduction rate, the length of generation, the intrinsic rate of natural increase and the stable age distribution. Despite a major attack in the 1940's on the net reproduction rate (as ordinarily calculated), and by implication on all other stable measures, this set of indices continues to thrive, perhaps *faute de mieux*.

The conventional net reproduction rate was questioned because it was female, period-oriented and age-specific. First, no formal basis exists for preference between a male and a female data base for the stable model, and the two outcomes generally differ. Efforts to force a determinate solution have been frustrated by the complexities of the marriage market. Because the likelihood of occurrence of a marriage in which the prospective spouses have any particular combination of characteristics is interdependent with the likelihood of occurrence for any other combination, the problem is essentially intractable, at least from the micro-analytic standpoint characteristic of other vital measures.

In the second place, the level and age distribution of net reproductivity in any period give a distorted impression of the behavior of constituent cohorts whenever there are intercohort variations in the reproductivity function. The assumption in the stable model of constancy for any period pattern of fertility and survivorship is tantamount to the assumption that cohort behavior will change—will accommodate itself to the prescribed period pattern. The criticism that stable measures, as conventionally calculated, are logically flawed can be met by changing habits, and basing stable models henceforth on cohort functions. To a practical man this is an unpalatable recommendation because contemporary cohort records are just not available in the same sense as those for periods. Furthermore, there is no generally satisfactory answer to the question of which cohort, of those currently fertile, should be the favored one. Of course the same might be said about period functions—why pick any one year rather than another?

—but the (unconsidered) preference is essentially always the most recent year. Were the cohort recommendation to be followed, the problems raised in the first section of this paper about how to complete truncated histories in some defensible manner would come to the forefront, not because they are defects of the cohort orientation, but because they force recognition of the embarrassing complexity of a changing reality. A final consequence of following the cohort recommendation would be that cohorts subsequent to the one chosen as base for the stable model would already be represented in the record, and probably by behavior that would give the lie to the assumption that the behavior of the base cohort should remain fixed.

The third charge against the net reproduction rate, *et al.,* is that there are as many sets of stable measures as there are sets of specified variables. Any assumption of fixity of rates, specific for N variables, implies coordinated variation in the rates specific for any N+1th variable. This damaging proposition concerning the empirical applicability of stable measures holds only when the period mode of temporal aggregation is used. The general process of specificity and control involves the derivation of exposure weights on the assumption of persistence of particular rates, and then their use as tenable replacements —tenable in the sense that, if the rates remain fixed, so will the weights —for the particular exposure weights observed for the population. Now the record for any cohort necessarily has exposure weights that make sense in relation to its rates, inasmuch as those rates are precisely what generate the cohort's evolving exposure distribution. It may be that heroic justification is required to use the idiosyncratic history of a cohort as the foundation for a model of eternal fixity, but at least the rates and the weights would be self-consistent.

Most of what we know about the relations among demographic processes and structures is the direct outcome of the study of the properties of the stable model, especially by Lotka. Nevertheless, the model and its derivative measures are limited in their applicability because the assumption of the fixity of vital processes is commonly, and certainly in the long run, implausible. The empirical record shows that the intrinsic rate of natural increase does not remain constant (except in the neighborhood of zero) for any more than a few decades. It is somewhat ironic that stable measures have provided ample documentation for the observation not that nonzero growth rates tend to persist (as they must in the stable model), but that they do not persist, either in the short run or in the long run. From general considerations it is evident that the average growth rate since the beginning of history must have been close to zero, and the average growth rate in the future will also be close to zero. More precisely, the past probably involved a sequence of growth rates occasionally departing from zero in a positive or negative direction; the future is likely to be the same, unless a persistent negative growth rate accomplishes our extinction.

One tradition in American population theory has practically died, and without mourners—that associated with the logistic. Little of value remains from the orgy of curve-fitting it inspired, but one important feature of the model deserves to be retained. In the logistic model, the growth rate increased from a zero asymptote for a while, and then declined to the zero asymptote again. This, of course, is the general character of the demographic transition, a period of disequilibrium between two kinds of stationarity. The interesting aspect of the logistic was its explanatory framework: the basic notions that the process of growth induces structural change (such as change in size), and that structural change induces change in the growth process. The stable model is an unrealistic description of how populations change through time because it contains no allowance for change of process in response to change of structure. The key theorem of the stable model is that structures are consequent upon processes, so that any initial structure is forced through time into conformity with the dictates of the fixed processes. This needs to be supplemented by a similar theorem, albeit a substantive rather than a deductive one, which runs in the opposite direction: whenever processes do not produce a stationary state, the consequent change in population structure will induce change in those processes in the direction of restoration of the stationary state.

This leads to the identification of an important future task: the design of models of demographic change in which the rate of growth is changing along a path that eventually leads to zero. This class of models may be called the ultimately stationary model. The likelihood is that the outcome will be less simple and less elegant than that of the stable model, because there may be many beginning points, and many paths to stationarity, and many types of stationarity. But at least computer simulation makes feasible the examination of an array of such options, and the derivation of quasi-mathematical solutions.

The distinction between cohorts and periods means nothing empirically in a model, like the stable, in which fertility and mortality are fixed, and therefore identical for periods and for cohorts. But in models in which processual change is tolerated, the cohort/period distinction becomes central to the characterization of population growth as process. Each cohort begins its existence at a certain size, which is the number of births in the time period identifying that cohort. The number of births is the product-sum of the period net reproductivity function (itself a translation of successive cohort net reproductivity functions) and the period age-distribution. In addition, the period age-distribution is a translation of successive cohort age-distributions. The last may be somewhat unfamiliar: each cohort can be thought of as a population, as a flow of person-years of life across the ages and times of the cohort's existence. The age distribution of the cohort as a population is in fact its survival function. The population with which we are most familiar (largely because of the way in which data are collected), is in fact a cross-section of these cohort age distributions; i.e., it is a translation of them in the same sense that period fertility is a translation of cohort fertility.

The idea of a cohort as a population leads to a revision of the notion of the net reproduction rate.[8] A cohort replaces itself not merely by bearing children, but by doing what is necessary to ensure survival of the children. Replacement is not a substitution of a number of births (in the child generation) for a number of births (in the parent generation), as would be implied by the conventional mode of calculation of the net reproduction rate, but rather the substitution of a flow of person-years in the child cohorts for a flow of person-years in the parent cohort. This variant would make no difference in any model in which mortality were presumed to be fixed, because each baby would then represent the same amount of life. But in a real context, replacement is more accurately determined by comparing the product of initial size and per capita person-years, for the children, with the product of initial size and per capita person-years, for the parents; i.e., $R_0{*}E''/E'$, where R_0 is the conventional net reproduction rate, and E' and E'' are the expectations of life of the parent and child cohorts respectively. In a more basic sense than that conveyed by the conventional net reproduction rate, the population is intrinsically growing whenever this product exceeds unity. This conceptualization of replacement in a person-years sense is not only more complete demographically than the conventional net reproduction rate, but is a more realistic image of the way in which a family or a society replaces itself, by an input of life to obtain an output of life.

An outline for an ultimately stationary model can be indicated.[5] Suppose an initial age distribution N(I, J) where I is age and J is period. If the population is assumed closed, this can be retrojected into its originating birth series, B(J–I) = N(I, J) /S(I, J–I), where S(I, J–I) is the proportion of the cohort born in year J–I that survives to age I. Then the future evolution of initial sizes of birth cohorts is given by

$$B(J) = \sum_{I=a}^{\beta} B(J-I) * R(I, J-I),$$ where R(I, J–I) is the net reproductivity

coefficient in age I for the cohort born in year J–I. (This is the finite form of a familiar basic equation of Lotka.) The formula is in effect a projection procedure that depends for its outcome on the evolution of the R(I, J–I) function. At the point where the net reproductivity coefficients reach their eventual stationary level, a simple formula indicates the ultimate birth cohort size:

$$B(W) = \sum_{J=0}^{\beta} (B(J-I) * (\sum_{I=J}^{\beta} R(I, J-I)) / \sum_{J=0}^{\beta} \sum_{I=J}^{\beta} R(I, J-I)$$

This leads immediately to an important parameter of the ultimately stationary model—the ultimate population size—determined by multiplying B(W) by the ultimate expectation of life at birth for the cohort. In most discussions of the consequences of demographic change, population size (typically varying with time) lurks ominously in the background, and serves less as a meaningful scientific variable than as a rhetorical device to celebrate the wonder of the exponential function. Perhaps ultimate population size may prove more useful than that.

In general, little attention has been paid to the kinds of demographic output that might be helpful for consequential analysis. Although much discussion centers on the importance of the age distribution, the only index in common use is the dependency ratio. For a population with mortality reasonably under control, there would seem to be advantages to the employment of indices of the shape of the B(J) function; i.e., the series of sizes of successive birth cohorts, considering that for most purposes the import of considerations of age distribution lies more in the comparative sizes of successive cohorts at the same age rather than of successive cohorts (ages) in the same period. Long-run variations in B(J) can provide a sense of the shape of the age distribution, and short-run variations capture the essence of irregularity of the age distribution.

Most analyses of the consequences of population change begin with a particular forecast, and then trace its consequences for one or another institutional sector, like education or the labor force. The likely inference from the exercise is that the future, sooner or later, looks grave, and something should be done about it—something that will nullify the forecast. What is missing in general is a specification of how grave that future is, and how much effort would be required to achieve a different future. More precisely, we need careful measurement to permit explicit weighing of alternatives. On the measurement side, we need to know the magnitude of changes required in various parameters of behavior to achieve one or another demographic future, and we need to know the magnitude of the changes in various outputs—those that presumably impinge on important societal functions. The form of such input and output measures is at present unclear, largely because the question has not been asked in this way. The final and more difficult task is to come to grips with what we know about the correlates of demographic inputs or, more to the point, the probably mutable correlates, to get an assessment of how much each possible future would cost, and then determine what we know about the correlates of demographic outputs, to get an assessment of how much it is worth to us to avoid each possible future. This is a considerable weight to place upon measurement and analysis, but it is what is needed to move from pure to applied demography.

MEASUREMENT AND ANALYSIS

The sophistication of fertility measurement, from a strictly methodologic standpoint, has increased impressively in the past four decades, and is now substantially in advance of the quality of the data with which we work. In turn, whatever the present defects of demographic information, in the double role of dependent and independent variables, they are clearly superior in clarity and precision to the qualitative and quantitative impressions that serve in lieu of measurements in the sociopsychologic, sociocultural and socioeconomic realms within which reproductive behavior is embedded. Methodologic progress has meant that we waste less time now trying to find explanations for changes that did not really happen, or at least did not happen in quite the way one might gather from a superficial glance. We can now ask rather well-formed questions, but our answers to these are not much better than what was available in the late 1930's to the group that gathered at Warren Thompson's to formulate hypotheses for the Indianapolis Study.

In such a quandary, one is sorely tempted to retreat within the shell of formal demography, where the world is neat and tidy and behaves in an orderly fashion (because that is the way we have constructed it), and leave to other social scientists the messy job of trying to cope with the flux of inchoate reality. We are properly proud of the rigor and discipline of our profession—and often unkind enough to show contempt for the flabby methodologic muscles of those who work beyond the demographic pale. In the writer's view, our relative success is explainable not so much by the rewards of virtue as it is by the circumstance that we have never let our reach exceed our grasp. We achieve satisfying answers because we ask easy questions.

Strictly from the standpoint of fertility measurement, causal and consequential, future progress seems guaranteed by our impressive stock of intellectual capital. The kinds of effort most needed now are likely to take us far afield from our comfortable methodologic niche, and force us into experience as fraught with the potentialities of humiliation as our record of fertility forecasts. But in the process, we may enlarge the scope of demography sufficiently to turn it into a genuinely rigorous branch of social science. More important, we may be able to assume that heavy but inescapable burden of beginning to meet some of our responsibilities for social policy.

REFERENCES

[1] Stolnitz, G. J. and Ryder, N. B., Recent discussion of the net reproduction rate, *Population Index* 14, 114–128, April, 1949.

[2] Whelpton, P. K., COHORT FERTILITY, Princeton, Princeton University Press, 1954.

[3] Whelpton, P. K. and Campbell, A. A., Fertility tables for birth cohorts of American women, Part 1, National Office of Vital Statistics, *Vital Statistics—Special Reports* 51(1), January 29, 1960.

[4] Ryder, N. B., The recent decline in the American birth rate, August, 1970, mimeographed.

[5] ———, Optimization of fertility policy, paper delivered at the annual meetings of the Population Association of America, April, 1971.

[6] Sheps, Mindel C., Simulation methods and the use of models in fertility analysis, Internation Union for the Scientific Study of Population, *Proceedings,* London, 1969.

[7] Ryder, N. B. and Westoff, C. F., REPRODUCTION IN THE UNITED STATES, 1965, Princeton, Princeton University Press, 1971.

[8] Ryder, N. B., Time as a demographic variable, paper delivered at the annual meetings of the American Association for the Advancement of Science, Chicago, 1959.

DISCUSSION

Philip M. Hauser: There are few examples in social science, including demography, in which research over the years has been as well designed to be additive in nature as the works of Norman Ryder. His innovative and persistent pursuit of a better understanding of fertility without question represents one of the major contributions to demography and, more specifically, to fertility analysis over the past two decades.

In his paper on Fertility Measurement, at least in first draft, and by Ryder's own admission, there are evidences of hasty preparation. In consequence, the paper may be described both as comprehensive and uneven. It is comprehensive in the sense that most of the important work that has been done in improving fertility measurements over the past several decades is at least mentioned. It is uneven in the sense that not all of the work is treated in uniform manner or intensity. This was perhaps to be expected. The most detailed treatment of advances in fertility measurement is given to that which Ryder, himself, has contributed. This lopsidedness, however, cannot be too much flawed because I think there would be general agreement that Ryder's contribution to improved measurement of fertility is, perhaps, the most important over the past forty years. To be sure, his work has been built upon foundations laid by others, including especially Whelpton in fertility and a number of others such as Lotka, Wicksell, Quensel, Vincent and Hyranius, in the development of the measurement of net reproduction and related rates. Ryder's unique contribution set forth in his paper is a definitive bridging of period and cohort fertility and the separation of factors of "quantity" and "tempo" as they affect both nuptiality and fertility.

Despite its general excellence a number of questions might be raised about the paper, at least in its initial form. Ryder, perhaps, makes too sharp a break between his "three distinctive analytic sectors" of fertility measurement because there is much in the way of interaction between "causal analysis" and "consequential analysis." Moreover there is a certain artificiality in depicting "the formal sector" as a link between the causal and consequential poles. In a broad sense "the formal sector" may be said to encompass both the causal and the consequential poles. Finally, it may be noted that Ryder has restricted his discussion almost entirely to consideration of what Hauser and Duncan, and

Lorimer before them, defined as "population analysis" as distinguished from "population studies."

That is, Ryder has confined his attention, and perhaps properly so, to consideration of fertility measurement concerned only with the interrelation of demographic variables considered either as independent or dependent. He has in the main ignored "population studies"—that is, efforts to get at the measurement of fertility that involves the relation between demographic and nondemographic variables. To be sure, there are not nearly as elegant models and procedures for measuring fertility in the realm of population studies as in the realm of demographic analysis. But it may be argued that the inability of the "formal sector," in Ryder's connotation of the term, to add more to the understanding of variations in fertility may in large measure be attributed to the lag in effective research in fertility in "population studies." More specifically, though much progress has been made in measuring the impact of nuptiality, timing, spacing, parity and age structure on fertility, we have not made much progress over the past forty years, at least as measured by definitive positive findings, on the impact of psychologic, social-psychologic, cultural and other factors in explaining fertility behavior.

A more comprehensive model than the one utilized by Ryder, which could serve as a paradigm for future advances in fertility research, would be one in which the "causal" and "consequential" sectors incorporated nondemographic as well as the demographic variables; and in which the "formal sector," as envisaged by Ryder, would include measurement methods effectively dealing with the nondemographic as well as the demographic variables. Ryder's observation that the causal sector has been too micro-analytic and the consequential sector, too macro-analytic up to this point may, in part, be corrected by the introduction of socioeconomic factors both on the macro and micro levels in each sector.

Undoubtedly, by reason of his restricted assignment, Ryder has made virtually no reference to the highly significant findings about fertility both in the causal and consequential sectors that have been developed over the years. In my judgment, it is correct to state that Ryder's substantive findings and his analysis of cohort fertility in the first part of the century in the United States has outmoded a large proportion of all that was thought to be known about fertility trends prior to his contributions.

In general, in his summary of methodologic advances Ryder, together with Nathan Keyfitz's recent presidential address to the Population Association of America, should provide demographers with a sound basis for satisfaction with what has been accomplished in the demographic analysis of fertility. Findings resulting from the methodologic advances have certainly added to the fund of knowledge, have penetrated public consciousness in a significant way, and have provided a foundation for the accelerating action programs devoted to the reduction of fertility and the control of rates of population growth.

Although the nature of data available from the census and vital registration systems was considerably improved to make possible the type of analysis Ryder presents, it is to be lamented that the quality of census and vital registration data has not kept pace with methodologic advances. In fact, as has often been pointed out, ingenious analytic techniques have resulted in large measure from the inadequacies of the data. But no matter how ingenious, methodology alone can never be an adequate substitute for better data. It is to be hoped that in the United States, as well as elsewhere, the need for more and much more refined basic statistics of the type needed because of methodologic advances will be heeded by government authorities whose business it is to collect the basic data.

Ryder's paper, in my judgment, is flawed by the relatively little attention it devotes to the problem of measurement as contrasted with analysis. Studies of response error affecting the reliability, validity and precision of information collected by census and survey methods has demonstrated the serious extent to which the data with which the demographer works are subject to error. Experience gained from the Current Population Reports of the U. S. Bureau of the Census, the statistics collected by the National Center of Health Statistics, the forthcoming report of the National Research Council on ways of dealing with the census errors, and part of the forthcoming mortality study by Kitagawa and Hauser on mortality, which employed matched records, point up the fact that much remains to be done to improve the quality of statistics with which the demographer must work. It is my judgment that advances in social science—generally, as well as demography specifically, are more retarded by inadequate measurement of the data at the collection stage than by deficiencies of conceptual frameworks or analytic methods. Demographers must continue to hammer away at the need for improving the quality of basic data.

Another respect in which Ryder's paper may be flawed is in its failure to at least mention the problems of measuring changes in fertility over short periods of time under the impact of family planning and other programs designed to reduce fertility. This is a most difficult problem on which much effort is being expended. It is conceivable that efforts to solve this problem may in many ways enhance the ability of demographers to measure fertility in other respects.

To become a better balanced paper, Ryder should elaborate on the contributions of others toward better fertility measurement, including the work of those whom he specifically mentions, including Coale, Zelnick, Brass, Sheps, Perrin, Potter, Tietze, Goodman and Keyfitz.

Finally, in closing, it is appropriate to call attention to some innovative ideas that almost always characterize a Ryder paper. Reference to "a remarkably simple formula" for determining "ultimate cohort size" and reference to cohort replacement involving not the ratio of number of children to the number of parents, but rather the ratio of the number of child person years to the potential parent person years merits further attention and utilization.

Without doubt, Ryder is quite correct in stating that the past forty years have put fertility measurement on a firm footing and that our methodologic future at least looks bright.

Wilson H. Grabill: Ryder mentioned that Pat Whelpton first organized his data into cohorts with the intention of getting bases for period rates. Whelpton, in many ways, was the "father" of cohort analysis in the US. I do not want to detract from that. However, Ryder is correct in saying that there were important predecessors. A few might regard Woofter's paper on generation reproduction rates as a forerunner of Whelpton's work. Maybe the French or others had still earlier work.