Note

1. In a loosely knit network. the individual spouse knows and has friendship ties with several people who do not know each other. In a tightly-knit network, there are mutual interactions among friends and spouse.

References

Bott, Elizabeth
1957 Family and social network. London, Tavistock.

Hill, Reuben
1967 The significance of the family in population research. *In* Family and fertility. William T. Liu, ed. Notre Dame, University of Notre Dame. Pp. 3–22.

Liu, William, Arthur Rubel, and Virginia Pato
1970 The Cebu family health project. Notre Dame, University of Notre Dame.

Nelson, Joel I.
1966 Clique contacts and family orientation. American Sociological Review 31(5): 663–72.

Rainwater, Lee
1965 Family design: marital sexuality, family size, and contraception. Chicago, Aldine.

Some Current Problems in Fertility Research: A Review

THOMAS W. PULLUM

During August 2–6, 1971, the International Association for Statistics in the Physical Sciences and the East-West Population Institute sponsored a conference at the East-West Center, Honolulu, Hawaii, titled "Statistical Problems in Population Research." This conference brought together 60 statisticians and demographers from around the world, representing government statistical offices as well as academic institutions, pragmatic as well as theoretical viewpoints, developed as well as developing countries. Sessions centered upon the presentation of some 18 papers and both prepared and spontaneous discussion of these papers.

During 1972 a volume containing these papers and prepared discussions will be published under the editorship of Professor Nathan Keyfitz (Department of Demography, The University of California, Berkeley). In the present article I shall attempt to summarize portions of the conference that may be of greatest interest to the readers of this journal. Persons desiring more complete discussion can refer to the forthcoming volume or to related items in Stinner's bibliographic article in this issue. In the present review I shall place less emphasis on the statistical nature of the problems than was present originally.

I shall focus on three major topics of the conference. The first of these concerns sociological determinants of fertility; the second concerns negative, as well as positive, medical aspects of birth control; and the third concerns the effectiveness of a family-planning program. Popular discussions of the population problem and family planning frequently gloss over these (and other) topics, which can be difficult to measure and to conceptualize satisfactorily, but which

are of central importance for safe and effective reduction of the growth rate.

Sociological Determinants of Fertility

Ryder. I shall consider two papers under this heading, one dealing with a developed country (the United States) and the other with a developing area (Asia). The first of these is "On sociological determinants of fertility in the United States," by Norman B. Ryder (Princeton University).

According to Ryder, the dominant theoretical model for explaining variations in fertility is based on the assumption of individual rationality. That is, people in a variety of circumstances set goals of family size which are somehow consonant with these circumstances, and then they act to implement these goals. There is the appearance of a duality, consisting of the goals and their implementation.

Unfortunately, of course, behavior is not so simple. The goals are altered by changes in circumstances of the couple, perhaps the most relevant of which is the birth of successive children. The model must permit the goals to be formulated in a sequential manner, taking into account the sex of children already born, changes in cultural norms, changes in family welfare, etc. The size and age-sex distribution of the completed family is then the result of a compounding of probabilities associated with each one of a sequence of decision steps.

Terms like "desired family size" oversimplify the "goal" aspect of the rationality model. Any such number must be subject to more or less modification, depending upon changing events.

Ryder argues that the forces which set and change fertility objectives are not so much a rational weighing of consequences by each couple involved as they are the norms of the couple's reference groups. Although this suggestion may sound obvious to social scientists, it is perhaps new for family-planning applications. Norms are usually "rational" only in the sense that they have enabled a society to maintain itself for a long period of time; they are not self-destructive of the society as a whole. A normative model and a rational model are thus basically different.

Two reasons are offered against the model of individual rationality. First, much reported fertility behavior is clearly irrational, in terms of matching of stated objectives with actual behavior. Fertility occurs when women (or couples) claim it is unwanted, even though the parties know of family-planning methods, can afford them, and express no objections to using them. Psychological and/or sociological factors evidently are involved.

Second, there is evidence that much behavior which is claimed as intentional has, in fact, been rationalized after the fact. It is well-known that the "desired" family size tends to increase as the completed size comes to equal or to exceed the "desired" size. Persons may refuse to state any goals at all in this area. Perhaps the correspondence between conception and sexual activity, which is surrounded with strong emotions and taboos, inhibits many persons from formulating goals which they should (for the sake of cognitive consistency) then effectuate.

Next consider the other aspect of the rationality model: the actual means of regulation. These logically fall into three categories: control of exposure, control of conception, and control of birth. The first of these is defined as most highly personal of the three, has been little studied, and presumably is not highly rational with respect to family planning. This area includes preferences and motives for marriage, divorce, etc. Control of birth (by induced abortion) has moral overtones and is not legal in most countries; moreover, for reasons of personal privacy it has seen little research.

Of the three means of regulation, control of conception has been the focus of interest. Choice of contraceptive is subject to sanction by religious groups and governments. Non-rational factors may prohibit the use of some specific, highly effective methods. For example, there may be resistance at a cultural level to ingesting un-

known substances, e.g., the Pill, despite all "scientific" information to the contrary.

Thus cultural norms play a large role in all phases of the family-planning process, although the rationality model does seem justified as a standard of reference for some survey work. Noting that there is much more variation in fertility behavior than can be explained by either model (there is great variation in behavior even within specific reference group structures), Ryder concludes that perhaps for poor and less-educated persons, in particular, planning – on any basis – is not a worthwhile activity. If one's world view is essentially fatalistic, then family planning may take neither rational nor non-rational dimensions; it may simply be irrelevant.

The preceding discussion pertains to the determinants of fertility in a general way. Ryder also describes some current group differentials within the United States. We briefly refer to some of his discussion, based on the 1965 National Fertility Study.

Roman Catholic fertility was about 30 per cent higher, per woman, than Protestant fertility, both as realized and as desired. Increased education lowered fertility for Protestants, but higher education in Catholic schools served to reinforce the desire for a large family, in contrast to secular higher education. It is clear that the primary source of higher Catholic fertility is the desire for more children, and not the failure or non-use of contraception.

Looking at another dichotomy, Blacks are more fertile than Whites, by about 20 per cent, but their *desire* is for slightly *fewer* children than are born to Whites. Moreover, the college-educated Blacks do manifest a lower fertility than the college-educated Whites. The reason for the high fertility of less-educated Blacks is that they are less likely to employ contraception, and the methods used are relatively inefficient.

The well-known inverse relation between fertility and socioeconomic status has, in fact, declined in intensity over the past few years. One explanation may be that formerly lower-status individuals had large families not because they desired them, but because contraception was unknown or unavailable to them. The difference may have resided largely in terms of implementation, rather than desired family size. Socioeconomic differentials in fertility may also be decreasing because above a certain security of income, resources may not be relevant to fertility.

Other differences: rural fertility is higher than urban fertility, probably a residual from the period in which large farm families were an economic asset. Increased education reduces fertility levels, perhaps, Ryder believes, largely because it improves the awareness of fertility regulation. The various variables cited are of course related; for example, rural areas are also areas of lower education.

Jain. Many of Ryder's observations extend to other countries. The second paper of this section is "Identification and measurement of socioeconomic factors affecting fertility in Asia," in which Anrudh K. Jain (Ford Foundation, New Delhi) conceptualizes fertility correlates and briefly describes empirical differentials in Taiwan, in particular. Jain participated in the extensive research in Taiwan that is also described elsewhere in this review article.

In nearly all fertility studies, fertility behavior is recorded up to the time of the study, spanning a period of many years, but socioeconomic characteristics are taken as of the date of the study itself. Some of these characteristics, which are assumed to be related to fertility, will not have changed (e.g., education and birthplace). But many others of importance (e.g., income and husband's occupation) will have altered, and their effect on fertility at a previous date will be indeterminate. It is possible that this lack of correspondence has caused many relationships to appear weaker than they should.

Jain suggests that socioeconomic variables be roughly classified into three groups: background variables, whose value precedes all fertility; current variables, which have changed in value over the years; and life-style and attitudinal variables. The last group consists mainly of value

orientations, measured by such items as newspaper readership, type of home furnishings, etc.

Research on Taiwan has shown that insofar as the above kinds of socioeconomic variables reflect a modern-vs.-traditional continuum, the most modern persons consistently have the lowest fertility. Income stands out as a variable which does not indicate the level of modernity, and it does not correlate well with fertility. Note that in more developed countries, income *does* correlate strongly (and negatively) with fertility.

As with the United States, the wife's and husband's levels of education correlate negatively (and about equally) with fertility. A rural background is associated with higher fertility, even among couples who have migrated to urban areas. The extended family is relatively common in Taiwan and is perhaps the best single indicator of traditionalism; the degree of family extension ranks with the educational level of the couple in explanatory power. The high-fertility family may thus be stereotyped as extended and rural, with both the husband and wife having low levels of education.

Intended and Unintended Effects of Fertility-Control Methods

Schrogie. The first paper to be described in this section is "Assessing adverse effects of the oral contraceptives: scope of the problem," by John J. Schrogie, Jr., M.D. (National Institute of Child Health and Human Development, Washington, D.C.). Public controversy in 1970 regarding side effects of the Pill led many women to discard its use for reasons that were often inconclusive. Dr. Schrogie's paper is an informed view of the hazard level associated with oral contraceptives and the problem of measuring this hazard.

There are many reasons why it is difficult to measure the Pill's adverse effects. First, the population of users is healthy, by and large, as well as young and geographically mobile, and is thereby unavailable for careful clinical examination. Oral contraceptives have been little used by organizations in developing countries, where follow-up might be easier. Typically, the decision to use oral contraceptives is made independently of any institutional family-planning program, and record-keeping is poor or non-existent.

After the problem of accessibility, a second difficulty is that experimental criteria are lacking. Selection of type of contraceptive or even type of Pill is not random, but is a complex (and unknown) function of other characteristics of the woman, which may themselves pre-dispose her to certain kinds of illness. Moreover, even when some randomness can be achieved in a supervised study, the woman will know whether her contraceptive is oral or not, as will most of the researchers and analysts, so that the expectations of all persons involved can bias the results.

A third difficulty, which would much hamper even a careful design, is the delay until presumed adverse effects appear. The latent period may range from a few months to many years, during which time all symptoms are normal.

Finally, all known adverse effects are rare. They occur as exaggerations (in frequency and/or severity) of illnesses which are rare in the absence of contraception. It is difficult to be statistically confident that a rare event among non-users is "less rare" among users, given the preceding complications of design, etc.

Dr. Schrogie emphasizes that the majority of side effects observed are those which can be expected from the chemical nature of the Pill. There is much variety in the composition of oral contraceptives but they generally are based on two hormonal compounds, an estrogen and a progestin, which occur normally in a woman as part of the menstrual cycle. Various illnesses associated with phases of the normal cycle may simply be emphasized by the intake of additional hormones, without being qualitatively different. If a woman has a history of such complaints, then she should be advised to use a different form of contraception or, at least, to use a Pill which has a favorable balance of the two basic hormones.

Some "common" examples of estrogen effects, for example, are fluid retention, nausea, headache, and excessive tiredness. During the later, progestin phase of the cycle, some com-

plaints are pre-menstrual depression, acne, appetite increase, and cramps. Adverse effects of the Pill, when they appear, will usually take the same basic forms. Dr. Schrogie emphasizes that therapy can be a simple change to a different brand of oral contraceptive. Certain early brands are no longer available because of their relative frequency of side effects.

The one unexpected and severe consequence is the increased chance of the formation of blood clots, with possible resultant heart failure or brain damage. This effect is "severe" in the sense that the outcome can be fatal, but the event remains statistically rare. In addition, those women who are susceptible to thromboembolism (blood clotting) in general will be those most likely to be affected by the Pill; as before, the Pill simply exaggerates a predisposition. High-risk women exhibit a number of factors, some of which are hypertension, a history of heart disease, arthritis, obesity, recent childbirth, cigarette smoking, diabetes, and later age.

In order to maintain a perspective, Table 1 lists some annual death rates per 100,000 women from a number of causes, including oral contraceptives. Dr. Schrogie drew this table from a 1968 article in the medical literature.

The table shows that Pill users have nearly eight times the chance of dying from thromboembolism than does the non-user population. Presumably, with time and experimentation this multiplier will be reduced in value. It is critical, however, that one balance this rate against the death rate associated with an unintended pregnancy, which may result from some alternative contraceptives or from the absence of contraception. The risk of mortality asscciated with pregnancy is very much greater than that associated with the Pill; unfortunately, most discussion has overlooked this standard of comparison.

Table 1

Estimates of risk of death from pulmonary embolism or cerebral thrombosis in users and non-users of oral contraceptives, compared with risk of death from certain other causes (after Table 4 of Dr. Schrogie's paper)

Population and cause of death	Age in years	
	20–34	35–44
Estimated annual death-rate/100,000 healthy, married, non-pregnant women from pulmonary or cerebral thromboembolism:		
Users of oral contraceptives	1.5	3.9
Non-users of oral contraceptives	0.2	0.5
Annual death-rate/100,000[a] total female population from:		
Cancer	13.7	70.1
Motor accidents	4.9	3.9
All causes	60.1	170.5
Death-rate/100,000[a] maternities from:		
Complications of pregnancy	7.5	13.8
Abortion	5.6	10.4
Complications of delivery	7.1	26.5
Complications of the puerperium:		
Phlebitis, thrombosis and embolism	1.3	2.3
Other complications	1.3	4.6
All risks of pregnancy, delivery and puerperium	22.8	57.6

[a]From the *Registrar General's* statistical reviews of *England and Wales,* for the year 1960.

To summarize, at the present time the measurement of hazard level is difficult and any numbers quoted are subject to doubt and qualification. But as nearly as possible, the use of oral contraception should be weighed against two factors: the woman's prior history of complaints that may be exacerbated, and the availability of other contraceptives, with their associated risks of side effects and of failure (i.e., of pregnancy).

Potter. The second paper to be considered in this section is "Additional births averted when post-conceptive methods are added to pre-conceptive methods of fertility control," by Robert G. Potter (Brown University). The "post-conceptive method" referred to in the title is, of course, induced abortion. Potter is well known for his research on births averted by IUDs and other methods of contraception, and this recent paper considers the reduction due to varying levels of abortion.

Potter uses the procedure of computer simulation, which has proven highly successful for analyzing complex patterns of behavior. Computer (or Monte Carlo) simulation involves specification by the researcher of a statistical "tree" of possible states of existence of hypothetical subjects, with the probabilities associated with each branch of this tree. For example, for a given hypothesized female, we may assign probabilities of marriage at age 15, 16, 17, etc.; then, once she has married, we may assign probabilities to her having a child by a certain date, etc. The computer is used to follow through a large group of these idealized subjects, aging them and individually tracing out their life patterns according to the specified probabilities. The researcher can "observe" the characteristics of this group at any point in time, totalling the number of first births by a given age of woman, etc. The utility of the method rests in the option of altering the specified probabilities, the parameters of the process. The effect of a later average age at marriage, etc., can then be detected readily.

Theoretically, any question that can be answered by simulation can also be answered by solving a corresponding set of mathematical formulas. In practice, however, formal solution can be much more complex. Simulation is being used increasingly in several areas of social science, particularly in Scandinavia and the United States.

Potter considers two groups of hypothetical women. The first group is subject to the events of pregnancy, miscarriage, stillbirth, and live-birth, with a post-partum period of sterility (amenorrhea) which varies with the age of the woman. In short, its fertility behavior corresponds closely in structure to that of an actual group of women. The parameters are approximately those of Taiwan, for which detailed data are available. In addition, the women in this group can be "treated" with contraceptives of varying levels of effectiveness, beginning at specifiable ages (age 30, for example).

The second hypothetical group corresponds exactly with the first except that a woman will always utilize induced abortion if her contraceptive fails. That is, abortion is used only as a backstop method, on those occasions when an inefficient contraceptive leads to the pregnancy of a woman who has accepted the principle of family planning. Abortion is not considered to be an alternative to pre-conceptive methods.

Due to contraceptive failure the first (control) group will have B births more than the second group; the second (experimental) group will have A abortions. The ratio B/A can be interpreted as "births averted per abortion." This ratio cannot exceed unity, since B and A both refer to the same set of unanticipated pregnancies and the event of birth is always a later pregnancy termination than the event of abortion. The reader may be surprised, however, to learn that B/A can be substantially less than unity. Potter obtains ratios as low as .44, such that more than two abortions are required to prevent just one birth (on the average).

The first reason why B/A can be much less than unity is that a pregnancy which is terminated artificially might have ended in a miscarriage or stillbirth anyway. The earlier the date of abortion, the greater the chance that a mis-

carriage would have normally occurred, i.e., the greater the chance of "redundancy."

The second factor is that the woman's sterile interval during a normal pregnancy and post-partum amenorrhea will be much reduced. She will be exposed to conception (or inefficient contraception) about one year earlier than if she proceeded with the birth. Thus, she may require two abortions in an interval of time that would otherwise have resulted in just one birth.

It may also seem paradoxical that abortion is most efficient, as a backstop method, when the pre-conceptive method is most efficient. (Although, of course, if the pre-conceptive method is perfectly efficient, then abortion is never called for). This phenomenon is tied to the reasoning of the preceding paragraph. If the pre-conceptive method is poor, then an abortion can be expected to be followed quickly by another pregnancy, and a larger number of abortions will eventually be required. On the other hand, if the contraceptive method is highly effective, then a long period of time is likely before the next abortion, and a smaller number of abortions will be required. In the process, the efficiency of each abortion is raised. The explanation can also take the following form: each abortion returns the woman to her pre-conceptive method; if this method is relatively efficient then, indirectly, the abortion will have been more efficient.

Certain policy implications are obvious. If a society chooses to utilize induced abortion as a supplementary method of family planning, then the efficiency (B/A) of each abortion can be increased in several ways. First, abortion should be used to supplement pre-conceptive methods that are themselves efficient. A woman who is using a generally effective method, even though it has failed, has also shown her greater desire for family limitation. Second, regardless of what (if any) pre-conceptive methods are in use, abortions to older women will be more efficient. The reduced fertility of older ages has the same effect as planned contraception, and also increases the efficiency of any contraceptive being used. Third, abortions should be performed as late as is medically and ethically advisable. About 60 per cent of all miscarriages occur during the first three months, and if the abortion is performed at the end of three months there is a much reduced chance that it will be redundant than if it is performed as soon as the pregnancy is recognized.

At present, of course, induced abortion is considered undesirable by many family-planning organizations, and it is illegal in many countries. Potter's paper provides more than a description of abortion efficiency, however. It illustrates an approach (computer simulation) which can project the consequences of any family-planning technique, given a range of assumptions, as well as many other processes of social and cultural change.

Evaluating the Effectiveness of a Family-Planning Program

Because of general agreement that most of the world has a population problem, a wide range of solutions has been offered. Some programs treat population control as a technological problem, a matter of making birth control materials available. Some programs integrate family planning with existing health care. A few programs attempt to discover and to alter the motives behind large families.

To evaluate these programs it is necessary to find the extent to which fertility has been affected, and at the same time to relate quantitatively the change in fertility to change in program inputs. Since demographic and socioeconomic factors quite distinct from the program may also have changed and thereby affected fertility, one must be able to measure and to interrelate a large number of variables.

The establishment of causal relationships between fertility and a variety of program and non-program-related variables is obviously both difficult and important. Comparative evaluation can tell us, for example, the relative cost, with different methods and programs, of each birth averted. One program may be best for a "resistant" set of socioeconomic variables and cultural values, whereas a different program is best under

"favorable" circumstances. Most evaluative research is highly quantitative, but it should be mentioned that programs are implemented according to non-economic criteria as well; the cost/benefit ratio is always subordinate to the political and cultural situation. Largely this is as it should be, since the "cheapest" methods of fertility control would be coercive, and should be rejected.

Three papers dealing with program effects will be described here. The first two of these are based on essentially the same data from Taiwan, which has undertaken intensive family-planning and demographic-recording activities during the past decade. The third paper is in itself a brief overview of program evaluation.

Hermalin. First consider "Taiwan: appraising the effect of a family planning program through an areal analysis," by Albert I. Hermalin (University of Michigan). The bulk of Hermalin's analysis is concerned with the years 1966–1969. Up to and during this interval major emphasis was on intrauterine devices (IUDs) and most recruiting was by field workers, rather than through advertising, for example, a more recent technique.

A few of the field workers are nurses who devote about half their time to recruiting IUD acceptors. The vast majority of the field workers are called "Pre-Pregnancy Health Workers" and spend all their time in recruitment. All field workers give potential acceptors a "coupon" which leads to a discount if an IUD is inserted by a participating doctor. The doctor sends the coupon to the family-planning organization, which uses the information it carries to allocate credit for the recruitment and to analyze the characteristics of the acceptor. Since 1965, there have been about 100,000 new insertions each year, the annual equivalent of 6 per cent of the married women aged 20–44.

Certainly, one would expect the birth rate to fall in the presence of such a program. But it is also true that Taiwan has been developing rapidly in socioeconomic terms, and historically this kind of development has been accompanied by a reduced birth rate, even in the absence of a program. We may then ask, "To what extent would the birth rate have fallen anyway? To what extent are the IUD insertions serving as substitutes for other methods which led to a low birth rate in such countries as France and Sweden? "

We first look at how one can respond to such questions and then at Hermalin's actual results. Hermalin chose an areal analysis in which the observational units were 331 provinces or local areas. These units were small enough that he could assume some internal homogeneity. The analysis is also cross-sectional in time; rather than following the characteristics of each unit for several years, the author made use of variation over units (provinces) at a single point in time. However, he introduced a lag, since the birth rate does not immediately reflect changes in other variables. The independent variables were taken as of about 1963, and the dependent variable as of 1966. The analysis was repeated for birthrates in 1967, 1968, and 1969, testing various (3, 4, 5, and 6 years) lengths of lag.

Figure 1 illustrates the postulated relationship among variables. An arrow from one variable to another indicates the direction of effect. The three boxes on the right give the dependent variables, and the one on the extreme right, age-specific fertility (for all women aged 20–24, say) is the variable of greatest interest. Hermalin separates this variable into its two components, the proportion married and the fertility of married women, since these two components have somewhat different sources themselves.

The remaining variables are independent in this analysis; no assumptions are made regarding their origin. They are included by Hermalin partly because of their availability, partly because of their intrinsic interest, and partly because they are indicators of variables which cannot be measured. Except for the IUD Acceptance Rates, they are all unrelated to the family-planning program. (In some programs, later age at marriage might also be a program objective, but in this program it is not.) A review of the effects follows.

It is *expected* that (1) in provinces where

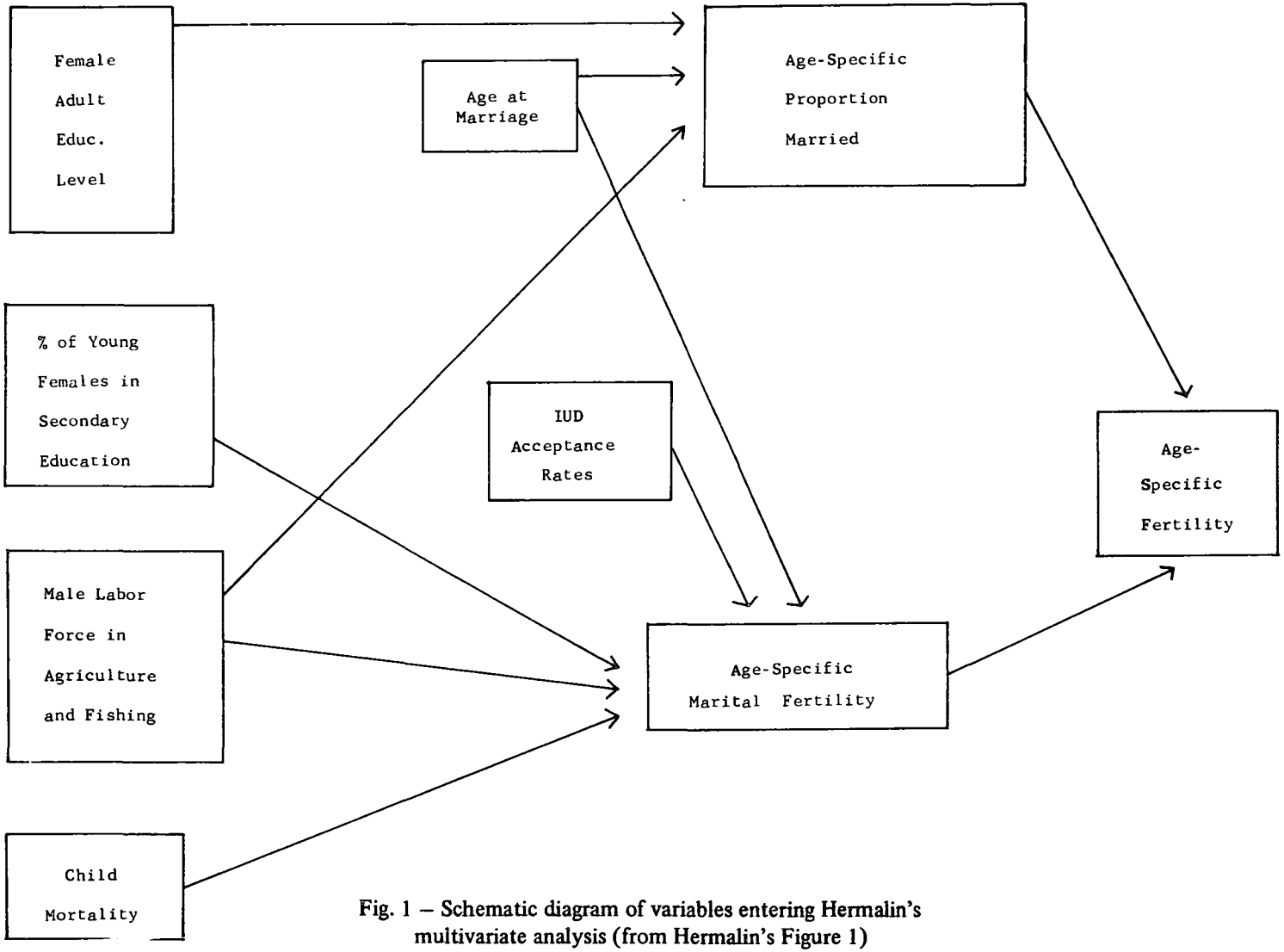

Fig. 1 – Schematic diagram of variables entering Hermalin's multivariate analysis (from Hermalin's Figure 1)

many women participate in adult-education programs, the proportion married will be lower; (2) if the regular education level is high (as measured by secondary-school enrollment) then marital fertility will be low (the "development" effect noted above); (3) the more agricultural (i.e., traditional) the province is, the greater the proportion marrying and the greater the fertility; (4) if child mortality is high, then fertility will also be high (as women try to replace children who have died).

Hermalin recognizes that some change in model is possible. For example, one might believe that the adult-education program would have a bearing on marital fertility and that general education might have an effect upon the proportion married. These effects would probably be weaker than those expected above, at any rate. There is also some element of arbitrariness in any choice of model.

Hermalin applies this model to four groups of women, of ages 20–24, 25–29, 30–34, 35–39. I shall not present here any of his numerical results (he used multiple regression) but shall describe his conclusions verbally.

(1) Adult education definitely does act to reduce the fertility rate, particularly for women under age 35 (who would be most involved by an adult-education program). Presumably this variable is an indicator of modernity and interest in planning for the future.

(2) The education level acts strongly to depress fertility, particularly in areas where fertility is already low. Here is evidence for increasing returns on education inputs after fertility has fallen.

(3) The prevalence of agriculture definitely is associated with high fertility, but primarily in that it increases the proportion who are married. It appears to have little effect on the fertility of married women, except that in areas which already have low fertility, agriculture *depresses* marital fertility even further. Hermalin does not offer an explanation; it is possible that this phenomenon is ecological and does not operate at an individual level. We note that low-fertility provinces have only a small proportion of farmers.

(4) The lagged child-mortality effect holds for older ages only. Women do not specifically try to replace children who have died except after age 30 or so.

Given a model that includes all of these variables, what is the effect of the program itself, as measured by the IUD-acceptance rate? The program had a highly significant impact on marital fertility, for all age groups, in the high-fertility provinces. However, its impact on areas of low and medium fertility was negligible except for women in their 20s.

In sum, the program was effective in high-fertility provinces, but change in low-fertility areas was primarily associated with increased levels of general education. The detailed analysis helps identify subgroups in which different kinds of inputs may be most effective. Hermalin concludes that IUDs are not simply substitutes for previous contraceptive techniques, at least in those areas of high program impact, but that the IUDs have reached a group of women with no previous method. On the other hand, in those areas where the program has had little impact on fertility despite an actual insertion program, the explanation is probably that IUDs have simply substituted for older methods.

Schultz. A second paper based on essentially the same data is entitled "A framework for analysis and its application to Taiwan's family planning program," by T. Paul Schultz (Rand Corporation).

There is good evidence that a major result of family-planning consciousness is a change in the spacing of children. In Taiwan the birthrate has dropped mainly for women over age 30, and in many areas women in their 20s are having *more* children. Schultz and others believe that when young, better educated women learn that highly effective contraceptive methods are available, they discard less effective methods and have their desired number of children while still young, thereby freeing themselves from childbearing and the care of young children during

their later years of full contraceptive use. The most diagnostic age-specific rates are therefore those of the women in their 30s, rather than their 20s.

Schultz's model has a different form and substance than Hermalin's but the two models together serve to illustrate the breadth of approaches to the same basic problem and data. The model is illustrated in Figure 2. The dependent variable is not a rate, but rather a numerator, i.e., the total number of children ever born to a woman of a specific age. The model is phrased in terms of the individual woman as unit, although the data forced Schultz to use areal units, as with Hermalin, for the numerical work.

The dependent variable is conceptualized to have three basic components. The first of these is the child-mortality rate experienced or expected by the mother. The second broad category is the number of children *desired* by the parents, which is not directly available for Taiwan but which is considered to be a function of environmental factors, such as the economic cost of the child, his expected economic benefit, etc. The third type of effect derives from local family-planning activities, which are assumed not to effect the desired family size but simply to facilitate its achievement. Figure 2 illustrates with shaded boxes the two aspects of the Taiwan program: the child-maternal health program, locally implemented by the health nurses mentioned above, and the family-planning program, based on the health workers also mentioned above. Recall that the nurses devote about half their time to family planning.

Schultz's analysis differs from Hermalin's mainly in his measurement of program input. He divides this into local man-hours of investment by the nurses, and man-hours of investment by the health workers. In the early years of the program, the health nurses had a far greater impact in reducing fertility, per man-hour of this activity. There is evidence that their effectiveness has remained high, but the program has concentrated on the other type of worker, so that by 1968 there were already 20 health workers for every nurse, and the ratio has continued to rise. Schultz suggests that the present mix of workers and nurses is not economically optimal. A much smaller number of nurses, engaged in a larger program of child and maternal health care, would accomplish as much as the large numbers of field workers, with their narrow range of activities.

A second conclusion regards the marginal effectiveness of additional workers. In the first years of the program, a fixed increment of man-hours would produce a much smaller reduction in fertility if an area had a large field force than if it had a small field force. This was an expected instance of the economic principle of diminishing returns; productivity per worker would be greater if workers were spread thinly. By 1968, however, the marginal-effectiveness curve had become much less steep and the size of an area's existing field force had less to do with the productivity of a fixed increment of man-hours. This change, perhaps a result of improved local organization and better deployment of field workers, would mean that the central organization should pay less attention to the size of local offices, relative to other criteria, in assigning new manpower.

Not only has the marginal productivity curve become less steep; it also indicates a reduction in the overall effectiveness of an increment of field time. By projecting productivity functions ahead in time, using Schultz's methods, one is able to estimate the date of worker saturation. At such a time an increment in manpower will have so small a return, in terms of reduced fertility, that it will be an unsound investment.

Ross. Finally, we consider a paper by John A. Ross of the Population Council, New York, "An overview of approaches to the estimation of program effects on fertility." Ross's paper does not go into detailed analysis of any particular program, but brings together the various problems and criteria associated with evaluation. Some of these have already been mentioned here and need not be repeated.

At the outset Ross emphasizes that evaluation must be relative to the program's goals. For ex-

ample, in Latin America, illegal and dangerous forms of induced abortion are a traditional method of fertility control. One objective of programs there is to draw women to safer methods, and therefore the substitution effect is more than tolerable; a program *can* be successful even if it has only altered the form, and not the extent, of fertility control.

A change in a birth rate may be divided into three logical components. One of these is the direct effect of the program; a second is the effect of uncontrollable changes, such as a changing age distribution or a change in non-program contraception; a third effect is the "interaction" of the preceding two. The interaction consists of the change brought about by the program *only in the presence* of the uncontrollable changes, or vice versa. Ross combines the first and third effects as the "total effect of the program, acting within the setting of specified background influences."

How can we measure the effect of uncontrolled events? How much change in a rate would we have expected if the program had not been introduced? (This question is not wholly answerable; for example, if a given program had not been introduced, then traditional methods might have expanded with more effectiveness, a possibility which cannot be assessed.)

The first approach described by Ross is areal analysis, as performed by Hermalin and Schultz. Analysis of a large number of small observational units permits the use of multiple regression to separate out non-program influences. In the case of Taiwan this technique was possible because of the organizational structure of the program, the excellent reporting system, and the availability of data on non-program variables. In most underdeveloped countries, fertility programs are not amenable to this form of analysis.

The second general approach is experimental design. One could, for example, take a list of 20 towns and *randomly* select 10 of them for participation in the program, leaving the other 10 untouched. The average change in the "untreated" towns would be the *expected* change in the "treated" towns. Unfortunately, it is difficult to prevent the control areas from being partially influenced. Experimental design is hampered by over-enthusiastic family planners, who refuse to leave areas of any size untouched for long, thereby making evaluation difficult and probably, in the long run, leading to considerable wasted effort.

A standard of comparison between organizations and methods exists when one is applied to a specified set of areas and another to a different set. However, inference about relative quality of the organizations and methods is possible only if the various areas are selected randomly.

A third method for measuring program and non-program effects is "matching." One can avoid the problem of random assignment of several units and manage with purposive selection of as few as two areas (or towns) if they are carefully matched in advance. For example, out of a list of 20 towns, one could find the two which corresponded most closely in relevant characteristics, such as ethnicity, socioeconomic level, age distribution, age at marriage, etc. The program would be applied to one of these but not to the other (if the pair are well matched then it will make no difference which enters the program). How the remaining towns are assigned is irrelevant, except insofar as they may contaminate the untreated town. After a lapse of time, the changes in fertility can be compared.

The last major comparative method is based on "pre-to-post-acceptance decline." This method is the weakest, since the standard of comparison is not itself permitted to change over time. One selects a group of women and measures their fertility performance in the five years prior to the program. Given constraints on the age distribution of these women, we might expect the same number of births during the five years following acceptance of some method. The difference will be "births averted." One cannot easily extend to the total female population any analysis which considers acceptors alone, however. Women who enter a program are likely to be much more motivated to reduce their fertility than are women who resist the program. This

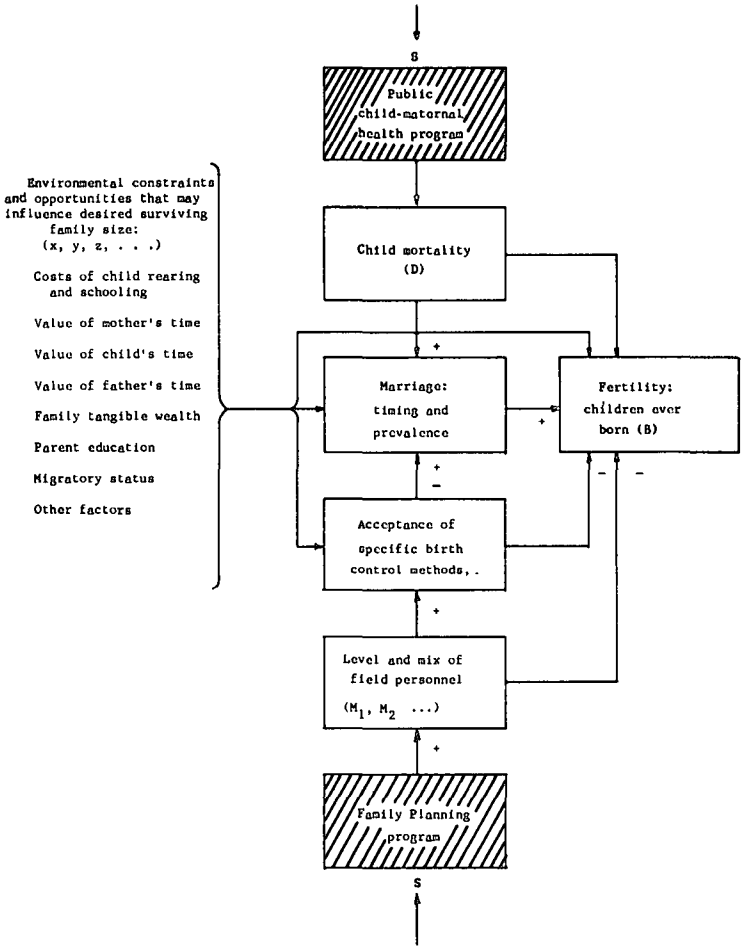

Fig. 2 – Schultz's model of fertility determinants (Figure 1 in Schultz's paper)

selection process will bias the results toward an exaggerated level of achievement.

Ross does not intend his overview to be complete, but it does subsume most evaluative methods. For example, the "matching" method could include the matching of one group of women with another on relevant characteristics, and then presenting birth control to one group only, or matching a group of adopting women (after adoption) with a group of non-adopters.

Conclusion

This review article has summarized a total of seven papers regarding basic issues in fertility and fertility control. The papers were grouped into three topics: first, the sources of fertility differentials, viewed both theoretically and empirically; second, technical and medical aspects of two methods for limiting births; third, some problems, criteria, and results in the area of program evaluation. Research in these areas is developing rapidly. This article was intended to make accessible those parts of a recent conference that may be most useful to the readers of this journal. The article has been expository, rather than critical, and the judgment of the author was exercised mainly in the selection of papers to be reviewed and in emphasis on certain portions of these papers.